

PF 02-APR-1987; 87JP-0076285.
 XX
 PR 02-APR-1986; 86JP-0076285.
 XX
 XX (RIKA) RIKAGAKU KENKYUSHO.
 PA
 XX WPI; 1987-325245/46.
 DR N-PSDB; AAN70651.
 XX
 PT DNA sequence coding cellulase gene - is derived from *Bacillus* sp. no.
 PT 1139z and can hydrolyse cello-triose and cello-tetraose.
 PS
 XX Disclosure; Fig 3-1 - 3-3; 11pp; Japanese.
 PS
 XX This cellulase hydrolyses cellobiose or cello-tetraose and does not
 CC hydrolyse cellobiose. Molecular weight is approx. 92k dalton and
 CC optimum pH is 9.
 CC
 XX
 SQ Sequence 800 AA;

Query Match 4.6%; Score 108.5; DB 8; Length 800;
 Best Local Similarity 19.9%; Pred. No. 0.26; Indels 173; Gaps 25;
 Matches 96; Conservative 54; Mismatches 159;

QY 8 LRLSLPCFALIAIOQAQVNPVAFVDEVRSEND---LGQDNELEPIDVQASATQASDTPT 64
 DB 10 LLSLIIIVLIIISL-----frcalaaegntrednfkhllgndnvkrpseagaldqdvvg 64
 QY 65 ANPL-DEHEPELY-----TTALENKTMLINCSAL-----NODIMRLACY-DTLVHGET 110
 DB 65 qmrlvdqhgkqkqlrgmsthgqlwfpelndnaykalandwesnmrlamygengyaen 124
 QY 111 PAVIKTKRSIRIDETI-----WQT-----IKGKPOVY 137
 DB 125 PELKSRVTKGIDLAENDMYIVDWHVHAPGPRDPRVYGAEDFFRIDIAALYPMNHII 184
 QY 138 YQETTDPR-----IFLMNGEKMLTKKDAKOLEYAKQFTPLSFSFDRN-----NPL 186
 DB 185 YELANEPSSNMNGGAGIPNNEEGWNAVK-----EYADPIVEMLRDSGNADHIIIVGSPN 239
 QY 187 WSSRP-----HNPMYVLPPIF--MHGKPNRS--PMTPSHEAKQFTPNFRAPDELKF 232
 DB 240 WSGRDLAENDPIDDHNTMYTVHFYGSNASTESYPPETPNSERGVMSNTRYA-----1 295
 QY 223 QVSIVKKAEDLMGT-----DSDLWFGYTOQ-----SHMQIFNGKNS-----R 270
 DB 236 ENGVAVFATE--WGSQANGDGARYFDEADVLEFINENIISWAMSLTN--KNEVSGAFT 352
 QY 271 PPRVH-----DYQPELFLNQPVYSDLPW--- 293
 DB 353 PFEIGKSNATSIDPPDQVWPEELISGEYVYRARIKYNYEP---IDTKYTKYVWDN 409
 QY 294 DCKVRMIGMGAHVHNSGES-----AKLSRSNRRAYLMAGMEKMLTYVPMRIW 340
 DB 410 dgltkgfgv-----ngdsrpedvvleneagalklsgldasndvsegunwanarisadgw 463
 QY 341 GR 342
 DB 464 gk 465

RESULT 15
 AAR26021

ID AAR26021 standard: Protein; 822 AA.

XX AAR26021;

XX 02-FEB-1993 (first entry)

XX Alkaline cellulase K-64.

XX Mass production; recombinant.

XX
 OS *Bacillus* sp. KSM-64 (FERM P-10482).
 XX
 PN JP04190793-A.
 XX
 PD 09-JUL-1992.
 XX
 PF 26-NOV-1990; 90JP-0324401.
 XX
 PR 26-NOV-1990; 90JP-0324401.
 XX
 PA (KAOS) KAO CORP.
 XX
 XX WPI; 1992-280112/34.
 DR N-PSDB; AAO27180.
 XX
 PT Alkaline cellulase mass prodn. - using recombinant plasmid and
 PT microorganism contg. cellulase gene
 PS
 XX Disclosure; Fig 4; 14pp; Japanese.
 PS
 XX Alkaline cellulase K-64 may be mass produced recombinantly in host
 CC microorganisms such as *E. coli* HB101, C600 or JM109, *B. subtilis* BD170,
 CC 168 or ISW1214.
 XX
 SQ Sequence 822 AA;

Query Match 4.6%; Score 108.5; DB 13; Length 822;
 Best Local Similarity 19.9%; Pred. No. 0.27;
 Matches 96; Conservative 54; Mismatches 159; Indels 173; Gaps 25;

QY 8 LRLSLPCFALIAIOQAQVNPVAFVDEVRSEND---LGQDNELEPIDVQASATQASDTPT 64
 DB 10 LLSLIIIVLIIISL-----frcalaaegntrednfkhllgndnvkrpseagaldqdvvg 64
 QY 65 ANPL-DEHEPELY-----TTALENKTMLINCSAL-----NODIMRLACY-DTLVHGET 110
 DB 65 qmrlvdqhgkqkqlrgmsthgqlwfpelndnaykalandwesnmrlamygengyaen 124
 QY 111 PAVIKTKRSIRIDETI-----WQT-----IKGKPOVY 137
 DB 125 PELKSRVTKGIDLAENDMYIVDWHVHAPGPRDPRVYGAEDFFRIDIAALYPMNHII 184
 QY 138 YQETTDPR-----IFLMNGEKMLTKKDAKOLEYAKQFTPLSFSFDRN-----NPL 186
 DB 185 YELANEPSSNMNGGAGIPNNEEGWNAVK-----EYADPIVEMLRDSGNADHIIIVGSPN 239
 QY 187 WSSRP-----HNPMYVLPPIF--MHGKPNRS--PMTPSHEAKQFTPNFRAPDELKF 232
 DB 240 WSGRDLAENDPIDDHNTMYTVHFYGSNASTESYPPETPNSERGVMSNTRYA-----1 295
 QY 223 QVSIVKKAEDLMGT-----DSDLWFGYTOQ-----SHMQIFNGKNS-----R 270
 DB 236 ENGVAVFATE--WGSQANGDGARYFDEADVLEFINENIISWAMSLTN--KNEVSGAFT 352
 QY 271 PPRVH-----DYQPELFLNQPVYSDLPW--- 293
 DB 353 PFEIGKSNATSIDPPDQVWPEELISGEYVYRARIKYNYEP---IDTKYTKYVWDN 409
 QY 294 DCKVRMIGMGAHVHNSGES-----AKLSRSNRRAYLMAGMEKMLTYVPMRIW 340
 DB 410 dgltkgfgv-----ngdsrpedvvleneagalklsgldasndvsegunwanarisadgw 463
 QY 341 GR 342
 DB 464 gk 465

Search completed: November 30, 2001, 14:17:02
 Job time: 384 sec

CC sequence of H. pylori (ATCC 55679) was determined from overlapping
 CC contigs generated by mechanically shearing the bacterial DNA. The
 CC sequences were analysed for ORF of at least 180 nucleotides, and the
 CC predicted coding regions defined by computer evaluation. To identify
 CC likely H. pylori antigens for vaccine development, the amino acid
 CC sequences predicted from various ORF were analysed for significant
 CC homology to other known or exported membrane proteins. Having identified
 CC and determined the sequences of interest, particular regions can be
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide
 CC production, e.g. in E. coli hosts.

XX
 CC
 SQ Sequence 253 AA;

Query Match 6.6%; Score 156.5; DB 18; Length 253;
 Best Local Similarity 25.6%; Pred. No. 1e-06; Mismatches 72; Indels 65; Gaps 8;
 Matches 56; Conservative 26;

QY 157 KDAKOLEVAAKQFTPLSLSDLRNTPLMSSRPHNPMYVLPFMHGKRPSPFESHE 216
 Db 72 kkylnmndylygtyflpfyhsf-----tpfgyyhninp----- 105
 QY 217 AKQFTPNEFAPRLKFOVSVKAAEDLWGTDSDFMGYQQSHWQJFNGKNSRPRVHD 276
 Db 106 ---ygrnef-----kfgisfrvprfihlwtkgtlyayqtawfglyndpqasapmrm 157
 QY 277 YQPEIFLTQVYSDLPMDGKV---RMIGMGAVHHSNG-ESAKLSRSNRNAVLAGMEMKN 332
 Db 158 fmpellyvyp1-nfkpryggkignfseilwqhlisngvgaqcyqpin----- 204
 QY 333 LTVMPRIWGRIFREGSGSQ--PDDNPDLIDYGYGDVRF 369
 Db 205 -----kegnpndqfpgqprvlykdyngqkdvtw 231

RESULT 13

AAW70991
 ID AAW70991 standard; Protein; 1686 AA.

XX AAW70991;

DT 19-OCT-1998 (first entry)

XX Human class II P13 kinase-C2alpha.

XX Human: class II phosphoinositide 1-phosphatase; P13 kinase;

KW PIK3-C2alpha; class II; resistance; Wortmannin; LY294002.

OS Homo sapiens.

XX MO9832864-AZ.

PN 30-JUL-1998.

XX 27-JAN-1998; 98WO-GB00244.

XX 28-JAN-1997; 97GB-0001652.

XX (LUDM-) LUDMIG INST CANCER RES.

XX Domin J, Waterfield MD;

XX WPI: 1998-427960/36.

XX N-PSDB; AAV42920.

XX New nucleic acid encoding phosphoinositol kinase 3-C2 alpha or its
 PT fragments - useful for, e.g. treatment of tumour cells where
 PT phenotype is associated with expression of kinase

XX Claim 3; Fig 1; 52pp; English.

XX The present sequence represents a human class II phosphoinositide 1-phosphatase
 CC (P13) kinase designated PIK3-C2alpha. It is characterised as a class II

CC kinase due to the presence of a conserved C2 domain found in murine and
 CC Drosophila class II P13 kinases, its apparent lack of a p85 binding site
 CC and a substrate affinity to inositol lipids ptdins and ptdins(4)P. The
 CC protein has resistance to P13 kinase inhibitors Wortmannin and LY294002.
 CC Antibodies against the protein (optionally humanised), are used to
 CC identify class II P13 kinases. Antisense sequences, antibodies or
 CC dominant negative mutants of the P13-C2alpha protein, are useful in human
 CC or veterinary medicine to block class II kinases. They can be used to
 CC treat tumour cells where the phenotype is associated with expression of
 CC P13-C2alpha protein.

XX
 CC
 SQ Sequence 1686 AA;

Query Match 5.0%; Score 117; DB 19; Length 1686;
 Best Local Similarity 19.2%; Pred. No. 0.13; Mismatches 154; Indels 140; Gaps 21;
 Matches 86; Conservative 67;

QY 45 QDNELPIDV-----QSATQASATPTANPLDE-----HEPELYTVALENTMLIN----- 88
 Db 522 eddetpvdlnkhlygiekpckeamttrhpveellasyhmqvelaqienqhravqylkav 1581
 QY 89 ---CSALNODIMRLACYDTLVHGETPAVYKTKRSIRLDETWTQTKGKPOVYQETDPI 1145
 Db 582 rkicsald-gyeltal-----tesvykkikravnlpr-----ktadv 618
 QY 146 FLWNGEMGLTKKAKOLEVAAKQFTPLSLSDLRNTPLMSSRPHNPMYV-----LP 199
 Db 619 slfge-----dtrsst-rgslnpenpvgslnqlta 651
 QY 200 IF-----MHGKPNRSPNTPSHEAKQFTPNEFAPRLKFOVSVKAAEDL---MCTDSPLW 252
 Db 652 lydlrlrhansgrprldeaqskvkeawttdqglfif-----aahglsnwnvsnyeky 707
 QY 253 FGTYQQSHWQJFNGKN-SRPF---RVHDIQPEIFLTQ-----PV-YSDLPMDGKYRMI 300
 Db 708 ylicslsh---ngkdlfkpryskkgvlyknfilylkwdellfpdlsqpldesvhl 763
 QY 301 GMGAVHHSNGSRSAKLSRSN-----RAYLAGMEMKNLTVMPRIWGRIFK 345
 Db 764 lfgyllngssgspsdnkqkgspealgkvsldplcdftrflltcg-----tkllylw----- 812
 QY 346 EGSSGPDNDPDIIDYGYGDVRFYQLEENKSNISGTVRNPBSGKALQDYYVPLGK 405
 Db 813 --tshsnsvpgrvttkgymerllyqdfpspafdllyttppqdrslilqnmlettend 870
 QY 406 ISGYFQIFQGYGQSLIDYNEATSFGV 432
 Db 871 lky-----klldlhkdsisgl 887

RESULT 14

AAW70420
 ID AAW70420 standard; Protein; 800 AA.

XX AAW70420;

DT 20-JAN-1991 (first entry)

XX Sequence encoded by cellulase gene derived from Bacillus sp. No. 1139.

XX Enzyme; cellotriose; cellotetrose; hydrolysis.

XX Bacillus sp. No. 1139.

XX Key Location/Qualifiers

XX Peptide 1..30

XX Protein 31..800

XX JPE2232386-A.

XX 12-OCT-1987.

Db 310 nysvslrlnpknktygataeaaytfpikgklkgvvrfgfygeslidyhknqnglgyimfn 369
 QY 438 DMMGL 442
 Db 370 dldgl 374

RESULT 9

AA70628
 ID AA70628 standard; Protein; 375 AA.

XX
 AC AA70628;

DT 18-JUL-2000 (first entry)

DE Neisseria meningitidis serogroup B strain ATCC13090 BASB033 protein.

XX BASB033; diagnosis; prophylaxis; treatment; antibacterial; vaccine;

KW Neisseria meningitidis infection.

OS Neisseria meningitidis.

XX WO200015801-A1.

PN 23-MAR-2000.

PF 09-SEP-1999; 99WO-EP06718.

PR 14-SEP-1998; 98GB-0020003.

PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

PI Ruelle J;

XX WPI; 2000-271439/23.

DR N-PSDB; AA252133.

XX Isolated BASB033 polypeptides and polynucleotides of Neisseria

PT meningitidis, useful for diagnosis, prophylaxis and treatment of N.

PT meningitidis infection -

PS Claim 4; Page 58; 93pp; English.

CC The present sequence is a BASB033 protein from

CC Neisseria meningitidis serogroup B strain ATCC13090. The protein

CC shows homology to the Klebsiella pneumoniae outer membrane

CC phospholipase A. The present sequence is useful for diagnosis,

CC prophylaxis and treatment of N. meningitidis infection. It may also be

CC used for the discovery and development of antibacterial compounds and

CC in vaccine compositions.

XX Sequence 375 AA;

SO

Query Match 34.4%; Score 812; DB 21; Length 375;

Best Local Similarity 44.4%; Pred. No. 1.5e-69;

Matches 162; Conservative 66; Mismatches 111; Indels 24; Gaps 8;

QY 87 INCSALMODIMRLACYTLVHGTPAVI-----KTKRSIRLDETWTQT-KGKPOVYVQE 140

Db 26 lqcaaldnvlrclycgrlfaagqpsaagegskavlnltelvrssldkgeavilvek 85

QY 141 TTDPFIFLMGNKGMILTKKDKOLEYAKQFTPLSLSDLDLNN-TPLMSSRPHPMYL 199

Db 86 ggdal-----pdsagetadylprslmydldkndrlgllgyvehmpmylmp 132

QY 200 IFMHGKPNRSPNPSH-EAKQFTNFRAPRLKFOVSQVKAADLMGTDLDLFGYTOO 258

Db 133 lwynspnydsplrtlttqekfsgqkreaetklyvafskiaedlfktradrldwfytyqr 192

QY 259 SHWQIFN-GKNSRPFRAVHDYQPELFTLPYVSDLPMDGKVMIMGAVHNSGSAKL 317

Db 193 sdwqlnygarksapfrntdykpeilflcpvkadlpfgrlrmjgafvhnsgnqsrpser 252
 QY 318 SMNRAVYLMAGMEKNTLVMPRIMGRIFKEGSGSPDDNPDLQYGYGDVRFYLOLENKS 377
 Db 253 swmllyamagmewgklvlpvwwvratdq-sgdk-nodpdldadygdydvklygrlndiq 310
 QY 378 NISGTVRNPNSCKGALQLDLYVYPLGKISGFQIFQSGSLIDYXHEATSPFCVGLMLN 437
 Db 311 nysvslrlnpknktygataeaaytfpikgklkgvvrfgfygeslidyhknqnglgyimfn 370
 QY 438 DMMGL 442
 Db 371 dldgl 375

RESULT 10

AAW98871
 ID AAW98871 standard; Protein; 355 AA.

XX
 AC AAW98871;

DT 31-MAR-1999 (first entry)

DE H. pylori GHP0 1723 protein.

XX GHP0 protein; Helicobacter infection; gastroduodenal disease; gastritis;

KW peptic ulcer disease.

OS Helicobacter pylori.

XX WO9843478-A1.

PN 08-OCR-1998.

PF 01-APR-1998; 98WO-US06371.

PR 29-JUL-1997; 97US-0902615.

PR 01-APR-1997; 97US-0833457.

PR 24-JUN-1997; 97US-0881227.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.

PI Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;

XX WPI; 1998-542293/46.

DR N-PSDB; AAX14390.

XX New Isolated Helicobacter polynucleotides - used to develop products

XX for the diagnosis, prevention and treatment of Helicobacter

XX infections and gastrointestinal diseases

PS Claim 8; Page 1976-1977; 2054pp; English.

CC This sequence represents a Helicobacter pylori GHP0 protein of the

CC invention. The polypeptides can be used for preventing or treating

CC Helicobacter infections, and gastroduodenal diseases associated with

CC these infections, including acute, chronic, and atrophic gastritis, and

CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be

CC used for the production of antibodies. The products can also be used for

CC detection and diagnosis.

XX Sequence 355 AA;

SO

Query Match 10.4%; Score 246.5; DB 19; Length 355;

Best Local Similarity 25.1%; Pred. No. 3.8e-15;

Matches 82; Conservative 41; Mismatches 117; Indels 87; Gaps 11;

QY 157 KKDAKOLEYAAKQFTPLSLSDLDLNNTPLMSSRPHPMYLVPFMHCKPNRSPNPSH 216

Db 69 kkyimmdyiglyflrpfyhsf-----tpdfgywhpnlp----- 102

01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.

(CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX

Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizzia R, Rappunli R, Ratli G, Scatato E, Scarselli M;
PI Tettein H, Ventler JC;
XX
XX WPI: 2000-062150/05.
DR N-PSDB; AAZ53920.
XX

Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics -
XX

Claim 2; Page 905; 1453pp; English.

AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54615 to AAZ5473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC manurestrial bacteria (e.g. meningitis and septicemia), to detect the
CC presence of *Neisseria* bacteria, or to raise antibodies. They may also
CC have been used to screen for agonists or antagonists, which may themselves
CC may also be used in gene therapy protocols.

Sequence 370 AA:

[illegible]

RESULT	8	
ID	AA70629	
XX	AA70629 standard; Protein: 374 AA.	
AC	AA70629;	
XX		
DT	18-JUL-2000 (first entry)	
XX		
DE	Neisseria meningitidis serogroup B strain H44/76 BASH033 protein.	
KM	BASH033; diagnosis, prophylaxis; treatment; antibacterial; vaccine;	
KW	Neisseria meningitidis infection.	
XX		
OS	Neisseria meningitidis.	
XX		
PN	WO200015801-A1.	
XX		
PD	23-MAR-2000.	
XX		
PF	09-SEP-1999; 99WO-EP06718.	
XX		
PR	14-SEP-1998; 98GB-0020003.	
XX		
PA	(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.	
XX		
PI	Ruelle J;	
XX		
DR	WPI: 2000-271439/73.	
XX	N-PSDB: AAZ52134.	
XX		
PT	Isolated BASH033 polypeptides and polynucleotides of Neisseria	
PT	meningitidis, useful for diagnosis, prophylaxis and treatment of N.	
PT	meningitidis infection -	
XX		
PS	Claim 4; Page 59; 93pp; English.	
XX		
CC	The present sequence is a BASH033 protein from	
CC	Neisseria meningitidis serogroup B strain H44/76. The protein	
CC	shows homology to the Klebsiella pneumoniae outer membrane	
CC	phospholipase A. The present sequence is useful for diagnosis,	
CC	prophylaxis and treatment of N. meningitidis infection. It may also be	
CC	used for the discovery and development of antibacterial compounds and	
CC	in vaccine compositions.	
XX		
SQ	Sequence 374 AA;	
Query Match	34.5%; Score 815; DB 21; Length 374;	
Best Local Similarity	44.7%; Pred. No. 7.6e-70;	
Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;		
QY	87 INCSALNDIMLACYDPLVHGERTAVI-----KKRSIRLDETIWQRI-KGKQVYQOE 140	
DB	25 IYcaaltldnvtllacydrdlfaaqlpsasagqegqeskavlnltetrssldkgyeavivex 84	
QY	141 TDPDFLEMGNEKGMGTLTKDKAKOLEYFAAKOFTPLSLSPFLDRNN-TPLWSSRPHNPMVLP 199	
DB	85 ggdal-----padsagetadityplslmyldkndrlgllgyvehnmpmylmp 131	
QY	200 IFMHGKPNRSPTPSH-EAKOFTPNFRAPAEIKFOVSVKVKAEDLWGTDSDLWFGYTOQ 256	
DB	132 Iwynpsnyapapsprgtctvggkfgqgkrraetklqvsfksklaedlfktradrldfwgylqr 191	
QY	259 SHMOJFN-GKNSRPRRVADYQPEFRLQPPVYSDLPMDCKVMIGMGAHHSNGSSAKLSR 317	
DB	192 scwqlyngagrsaprintcdykpelltlpvpvradlpfpggrlmltqagvltvngsqsrpestr 251	
QY	318 SNNRRAYLWAGMEKMKLITVMPRIGRIFFEFGSGSQPDNDPDLIDYVGGVDFVLQLENKS 377	
DB	252 swnrllymagnewgvlitlprvvratfdq-sgdk-ndnppdiadymygvdvklgylndrdq 309	
QY	378 NISGIVRINPNRSGKALDLDVYPLGKGKISGTFQIFQSGSLIDYVNHGATSPGVGLMLN 437	

PT Isolated BASB033 polypeptides and polynucleotides of *Neisseria*
PT meningitidis, useful for diagnosis, prophylaxis and treatment of *N.*
PT meningitidis infection -
XX
XX
PS
PR Claim 4; Page 59; 93pp; English.
CC
CC The present sequence is a BASB033 protein from
CC *Neisseria meningitidis* serogroup B strain H4476. The protein
CC shows homology to the *Klebsiella pneumoniae* outer membrane
CC phospholipase A. The present sequence is useful for diagnosis,
CC prophylaxis and treatment of *N. meningitidis* infection. It may also be
CC used for the discovery and development of antibacterial compounds and
CC in vaccine compositions.

```
Query Match Similarity    34.5%; Score 815; DB 21; Length 374;
Best Local Similarity     44.7%; Pred. No. 7.6e-70;
Matches 163; Conservative   67; Mismatches 111; Indels   24; Gaps      8;
```

OY INCSALNODIMLACVDPELVHGETPAVI-----KKRSIRLDLETWQT-KGKPOVVOE 140
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Db lqcaaltnvttlaeydrtlfaaqlpsasggqegekavlnlterysldkgeaviylpe 84

OY 141 TDPPIELMGNEKGMLTKRKDAQOLEYAARKOFTPLSLSEFLDRNN-TPLMSSRPHNPMYLP 199
|TDTI-----K----|::|::|::|::|::|::|::|::|::|::|::|::|
Db ggdal-----padsgagetadlyrplsimyldknrdrgllayvrehmpmylmp 131

OY 200 IFMHCKPKNRSPPTPSH-EAKQFTPNMFAPRLFKOVSYKVKAADLMOTDSDDLYFGCTYOQ 258~
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Db lwynsnpyapbspertgetvgsekfgqqkrtaeklkqvskfskliaellfkftradiyfgyqr 191

OY 259 SHMOJFN-GKNSRRPRVHDYOPEITLFOPPVSDLPMDCKVMIMGAVHNNGESAKLSR 317
+|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Db sbwqlyngyrksaprntcdykprrellcpvrkadrlpfgrlrilmiaagftnhsngsstriest 251

OY 318 SMMRAAYLAGMEKMKLVTMPRIWGRTFEKSQSOPDNDPDILDYGYGDVRFLYLENKS 377
|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Db swnrilyamagnegwglvtlviprvvrafdg-sgdk-nndnpdiadymgvydvkrylindrq 309

OY 378 NISGVTRNPBSRGKALOLDVVYPLGDKISTGFOLFPGSTGGSLIDYNHEATSPVGGLMN 437

```

Db      21  lqcaaltdnvtlacydriflaaqlpsaagqegqeskavlnltetvrsldkgaevlvvek 80
OY      141  TTDPFILMGNEKGM/LTKKDAKOLEYAAKOFTPPLSLSFDDLBN--TPLWSSRPHNPMVLP 199
Db      81  ggda1-----padsagetadlycpislmtydldkndrlgllgyrehnmpmlymp 127
OY      200  IFMHGKPNRSPNTPSH-EAKQFTPNFERAPBLKFOYSVKVKAEDLWGTSDMLFGYTOQ 258
Db      128  fwynspnyapsptgtvtgqefggkkractklyqsfkxaeenlktadtlwfygtqr 187
OY      259  SHMOIFN-GKNSRPFRRNDYQPELFLTPQVYSDLPMDGKVMGMGAVHNSGSAALS 317
Db      188  sdwqlyngqrksapfrntdykpeflftlpvkadlprfggrlrmlygaftvhsngsqrspre 247
OY      318  SMNRAYLMAGMEKNLTVMPRIWGRIKFEKSGSQPDNDPILDYGYGDVRFYOLENKS 377
Db      248  swmrlyamagmewgklvlprrvwtatfdq--sgdk--ndnpdlaadymgygdvklqyrlndrq 305
OY      378  NISGTVRYNPRSGKALQLDYVYPLGKISGYFOIFOGYQSLIDYNHAEATSFVGMLN 437
Db      306  nvysvlyrnpktygalaayaftfplkgklkgvvrfgfhgysldynhknqngiglmfn 365
OY      438  DMWGL 442
Db      366  dwdgl 370

RESULT 6
AA775157
ID  AA775157 standard; Protein; 370 AA.
AC  AA775157;
DT  21-MAR-2000 (first entry)
DE  Neisseria meningitidis ORF 582 protein sequence SEQ ID NO:1788.
XX
DE  Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW  antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW  antibacterial; gene therapy.
XX
OS  Neisseria meningitidis.
XX
PN  WO957280-A2.
XX
PD  11-NOV-1999.
XX
PF  30-APR-1999; 99WO-US09346.
XX
PR  01-MAY-1998; 98US-0083758.
PR  31-JUL-1998; 98US-0094869.
PR  02-SEP-1998; 98US-0098994.
PR  02-SEP-1998; 98US-0099062.
PR  09-OCT-1998; 98US-0103749.
PR  09-OCT-1998; 98US-0103794.
PR  09-OCT-1998; 98US-0103796.
PR  25-FEB-1999; 99US-0121528.
XX
PA  (CHIR ) CHIRON CORP.
PA  (GENO-) INST GENOMIC RES.
XX
PI  Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI  Petersen J, Piza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI  Tettelin H, Venter JC;
XX
DR  WPI: 2000-062150/05.
DR  N-PSDB; AA253919.
XX
PT  Novel Neisserial polypeptides predicted to be useful antigens for
XX  vaccines and diagnostics
XX
PS  Claim 2; Page 904; 1453pp; English.

```

```

XX      XX  AA253015 to AA254536, AA254577 to AA254615, and AA74253 to AA775941
CC      CC  represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
CC      CC  and polypeptides. AA254537 to AA254576 and AA254616 to AA254731 represent
CC      CC  PCR primers used in the exemplification of the present invention. The
CC      CC  polypeptides, the polynucleotides, antibodies and compositions of
CC      CC  the invention can be used as vaccines, as diagnostic reagents, and as
CC      CC  immunogenic compositions. The polypeptides can be used in the
CC      CC  manufacture of medicaments for treating or preventing infection due to
CC      CC  Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC      CC  presence of Neisseria bacteria, or to raise antibodies. They may also
CC      CC  be used to screen for agonists or antagonists, which may themselves
CC      CC  have use as antibacterial agents. The polynucleotides of the invention
XX      XX  may also be used in gene therapy protocols.
XX      XX
SO      SO  Sequence 370 AA;

Query Match 34.5%; Score 815; DB 21; Length 370;
Best Local Similarity 44.7%; Pred. No. 7.5e-70;
Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;

OY      OY  87  INCNALMODIMRLACYDTLVHGETRPANI-----KTKRSIRLDETINQTI-KGRQVYYQE 140
Db      Db  21  lqcaaltdnvtlacydriflaaqlpsaagqegqeskavlnltetvrsldkgaevlvvek 80
OY      OY  141  TTDPFILMGNEKGM/LTKKDAKOLEYAAKOFTPPLSLSFDDLBN--TPLWSSRPHNPMVLP 199
Db      Db  81  ggda1-----padsagetadlycpislmtydldkndrlgllgyrehnmpmlymp 127
OY      OY  200  IFMHGKPNRSPNTPSH-EAKQFTPNFERAPBLKFOYSVKVKAEDLWGTSDMLFGYTOQ 258
Db      Db  128  fwynspnyapsptgtvtgqefggkkractklyqsfkxaeenlktadtlwfygtqr 187
OY      OY  259  SHMOIFN-GKNSRPFRRNDYQPELFLTPQVYSDLPMDGKVMGMGAVHNSGSAALS 317
Db      Db  188  sdwqlyngqrksapfrntdykpeflftlpvkadlprfggrlrmlygaftvhsngsqrspre 247
OY      OY  318  SMNRAYLMAGMEKNLTVMPRIWGRIKFEKSGSQPDNDPILDYGYGDVRFYOLENKS 377
Db      Db  248  swmrlyamagmewgklvlprrvwtatfdq--sgdk--ndnpdlaadymgygdvklqyrlndrq 305
OY      OY  378  NISGTVRYNPRSGKALQLDYVYPLGKISGYFOIFOGYQSLIDYNHAEATSFVGMLN 437
Db      Db  306  nvysvlyrnpktygalaayaftfplkgklkgvvrfgfhgysldynhknqngiglmfn 365
OY      OY  438  DMWGL 442
Db      Db  366  dwdgl 370

RESULT 7
AA775158
ID  AA775158 standard; Protein; 370 AA.
AC  AA775158;
DT  21-MAR-2000 (first entry)
DE  Neisseria meningitidis ORF 582 protein sequence SEQ ID NO:1790.
XX
DE  Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW  antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW  antibacterial; gene therapy.
XX
OS  Neisseria meningitidis.
XX
PN  WO957280-A2.
XX
PD  11-NOV-1999.
XX
PF  30-APR-1999; 99WO-US09346.
XX

```


QY 181 RNNTPLMSSRPHPNMYVLPFIMHGKPNRSPNTPSHEAKQFTPNEFRAPELKFOYSVKVA 240
 CC and to identify agonists and antagonists. The polypeptides, antibodies,
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.

QY 241 AEDLMGTDSLWLMGYTQOOSHWQJFNGKNSRPFVHDYQPEFLFQPYSLPMDGKVRMT 300
 CC
 Db 241 aedlwgtcdslwfygtqgshtwqjfngrfvdhdyqpeflftqpysslpmdgkvrmt 300
 CC
 QY 301 GMGAVHHSNGESAKLSRSNMRAYLMAGMEWKNLTVMRIGRIFKEGSGSQPDNDPILD 360
 CC
 Db 301 gmgavhhsngesaklsrswmraylmagmeknltvmrigrifkegsgsqpdndpild 360
 CC
 QY 361 YYGXGQVRFLEYOLENKSNIQSVRYNPRSGKALQLDVYVPLGKISGYPQIFQGYGQSL 420
 CC
 Db 361 yygvgdvrflfyqlenksnlsqsvrynprrsgkalgldvvyplgkysgyfqiifggygs 420
 CC
 QY 421 IDYNHEATSFYGVGLMDNMWGL 442
 CC
 Db 421 idynheatsfygvlmdnmwgl 442
 CC

RESULT 3
 AAY85270
 ID AAY85270 standard; Protein: 442 AA.
 AC AAY85270;
 XX
 XX 29-JUN-2000 (first entry)
 DT
 XX
 XX BASB034 amino acid sequence #3.
 DE
 XX
 XX Moraxella catarhalis infection; BASB034; diagnosis; staging;
 KM vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KM sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KM hearing loss; antibacterial drug.
 XX
 XX Moraxella catarhalis.
 OS
 PN MO200015802-A1.
 XX
 XX 23-MAR-2000.
 PD
 XX
 XX 14-SEP-1999; 99WO-EP06781.
 PF
 XX
 XX 14-SEP-1999; 98GB-0020002.
 PR
 XX
 XX (SMRK) SMITHKLINE BEECHAM BIOLOGICALS.
 PA
 XX
 XX Ruelle J;
 PI
 XX
 XX WPI: 2000-271440/23.
 DR
 XX
 XX N-PSDB: AAA10702.
 PT
 XX
 XX Novel BASB034 polynucleotides and polypeptides from Moraxella
 PT catarhalis used to prepare vaccines against bacterial infections
 XX
 PS Claim 3: Page 68; 106pp; English.

CC and to identify agonists and antagonists. The polypeptides, antibodies,
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.

QY 1 MKVSLSTLTLSLPCFAILAIQQAAPNPVAFVDEVRENDLGQDNELPIDVQATQSA 60
 CC
 Db 1 mkvslstltlsiscfalaiaiqgakvpnpvafvdevrsendlgqdnelpidvqatqsa 60
 CC
 QY 61 STDANPLDDEHEPRLTYTALENKTMNLNCSALNODIMRLACYDTLVGEPVATKTRST 120
 CC
 Db 61 stdanpldehepelytalaenktmlncsalnqdmrlacydltlvgepvatktkrs 120
 CC
 QY 121 RLDETIWQTIKGRPOVYVOETTDPIFLMGNEKGMLTRKDAKOLEYAKQFTPLSLSPDL 180
 CC
 Db 121 rldetiwtikgrpqvvyvoettdpiflmgnekgmtlrkdakoleyaakqftplsld 180
 CC
 QY 181 RNNTPLMSSRPHPNMYVLPFIMHGKPNRSPNTPSHEAKQFTPNEFRAPELKFOYSVKVA 240
 CC
 Db 181 rnntpmwssrphnmyvlpfimgkpnrsptpsheakqftpnefrapelkfgysvkva 240
 CC
 QY 241 AEDLMGTDSLWLMGYTQOOSHWQJFNGKNSRPFVHDYQPEFLFQPYSLPMDGKVRMT 300
 CC
 Db 241 aedlwgtcdslwfygtqgshtwqjfngrfvdhdyqpeflftqpysslpmdgkvrmt 300
 CC
 QY 301 GMGAVHHSNGESAKLSRSNMRAYLMAGMEWKNLTVMRIGRIFKEGSGSQPDNDPILD 360
 CC
 Db 301 gmgavhhsngesaklsrswmraylmagmeknltvmrigrifkegsgsqpdndpild 360
 CC
 QY 361 YYGXGQVRFLEYOLENKSNIQSVRYNPRSGKALQLDVYVPLGKISGYPQIFQGYGQSL 420
 CC
 Db 361 yygvgdvrflfyqlenksnlsqsvrynprrsgkalgldvvyplgkysgyfqiifggygs 420
 CC
 QY 421 IDYNHEATSFYGVGLMDNMWGL 442
 CC
 Db 421 idynheatsfygvlmdnmwgl 442
 CC

RESULT 4
 AAY85269
 ID AAY85269 standard; Protein: 442 AA.
 AC AAY85269;
 XX
 XX 29-JUN-2000 (first entry)
 DT
 XX
 XX BASB034 amino acid sequence #2.
 DE
 XX
 XX Moraxella catarhalis infection; BASB034; diagnosis; staging;
 KM vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KM sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KM hearing loss; antibacterial drug.
 XX
 XX Moraxella catarhalis.
 OS
 PN MO200015802-A1.
 XX
 XX 23-MAR-2000.
 PD
 XX
 XX 14-SEP-1999; 99WO-EP06781.
 PF
 XX

XX Claim 3; Page 69; 106pp; English.

XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
PS strain Mc2969. The invention relates to BASB034 polypeptides from
CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
CC polynucleotides and polypeptides may be employed as research reagents and
CC material for the discovery of treatments and diagnostics for diseases,
CC particularly human diseases. They are particularly used to diagnose and
CC treat M. catarrhalis infections. They can be used for diagnosis of
CC disease, staging of disease, or determining response of an infectious
CC organism to drugs. The polynucleotides may be used as a source for
CC hybridization probes, and for screening of genetic mutations, serotype,
CC organism or strain identification, identification of mutations in BASB034
CC sequences, and as components of arrays which are useful for diagnostic
CC and prognostic purposes. The polypeptides can be used to produce
CC antibodies. The polypeptides can also be used in vaccine formulations,
CC and to identify agonists and antagonists. The polypeptides, antibodies,
CC agonists and antagonists (which are bacteriostatic) are used for the
CC treatment and prevention of diseases such as otitis media in infants and
CC children, pneumonia in elderly, sinusitis, nosocomial infections and
CC invasive diseases, and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.

XX Sequence 442 AA:

Query Match 100.0%; Score 2363; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 9,3e-219; Indels 0; Gaps 0;
Matches 442; Conservative 0; Mismatches 0;

QY 1 MKVSLSTLSTLPCFAITAIQQAQAVPVAFVDEVSENDLGDNDLPIDVOSATOSA 60
DB 1 mkvslstlslstlpcfallalqgaqavpvaafvdevsendlgdndelpdvgsatgsa 60
QY 61 STDTPNPIDEHEPELYTTALENKMILNCSANODIMRLACDYLTVHGETPAVITKRSI 120
DB 61 stdtlnpidehepeilyttalenkmlncsainodimrlacdyltvhgetpaviktrsi 120
QY 121 RLDETIWOTIKGKPOVVOYQETTDPIFLMGNEKGMLTRKDAKOLEYAAKOFPLSLSPDLD 180
DB 121 rldetiwtikgkpvovyoettdpiflmgnekgmtlrrkdakoleyaakofplslsfdld 180
QY 121 RLDETIWOTIKGKPOVVOYQETTDPIFLMGNEKGMLTRKDAKOLEYAAKOFPLSLSPDLD 180
DB 121 rldetiwtikgkpvovyoettdpiflmgnekgmtlrrkdakoleyaakofplslsfdld 180
QY 181 RNNPFLMSRPNPNVYLPPIEMHGKPNNSPNTPSHEAKOFTPNFRAPBLAKQVSVKKA 240
DB 181 rnnpflmsrpnnpnyvlppliemhgkpnnsntpsheakoftpnfrapblakqvsvkka 240
QY 241 AEDLGGTSDLMFGYTOQSHMOIFNGKNSRPRVHDYQEPFLQPVYSDLPMDCKVRMI 300
DB 241 aedlvgtsdldwfytcqshwqifngknsrprvhdypelldqpyvsdlpmdckvrtmi 300
QY 301 GMGAVHNSGESAKLSRSMNRAYLMAGMEKMLTVMPRIWGRIFKEGSGSQPDNDPDIID 360
DB 301 gmgavhngesaklsrsmnraylmagmekmltvmpriwgrifkegsgsqdpndpdiid 360
QY 361 YYGVDVRFVLQLENKSNISGTVRYNPRSGALQDLVYVPLGKISGYFOIFGYGOSL 420
DB 361 yygvdvrfvlyqlenknsnsgtvrynprrsgalqldvyyvplgkisygyfifgygysl 420
QY 421 IDYNHEATSPFGVGLMNDMMGL 442
DB 421 idynheatsfgvglmndmmgl 442

RESULT 2
AA85268
ID AAY85268 standard; Protein: 442 AA.
XX
AC AAY85268;
XX

DT 29-JUN-2000 (first entry)

XX BASB034 amino acid sequence #1.
DE
XX
XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW hearing loss; antibacterial drug.
XX
XX Moraxella catarrhalis.
OS
XX WO200015802-A1.
PN
XX 23-MAR-2000.
XX
XX 14-SEP-1999; 99WO-EP06781.
PF
XX
XX 14-SEP-1998; 98GB-0020002.
PR
XX
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
PA
XX
XX Ruelle J;
PI
XX
XX WPI: 2000-271440/23.
DR
XX
XX N-PSDB; AAA10700.
DR
XX

PT Novel BASB034 polynucleotides and polypeptides from Moraxella
PS catarrhalis used to prepare vaccines against bacterial infections -
XX
XX Claim 3; Fig 2; 106pp; English.

XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2931 (ATCC 43617). The invention relates to BASB034 polypeptides
CC from M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The
CC BASB034 polynucleotides and polypeptides may be employed as research
CC reagents and material for the discovery of treatments and diagnostics for
CC diseases, particularly human diseases. They are particularly used to
CC diagnose and treat M. catarrhalis infections. They can be used for
CC diagnosis of disease, staging of disease, or determining response of an
CC infectious organism to drugs. The polynucleotides may be used as a source
CC for hybridization probes, and for screening of genetic mutations,
CC serotype, organism or strain identification, identification of mutations
CC in BASB034 sequences, and as components of arrays which are useful for
CC diagnostic and prognostic purposes. The polypeptides can be used to
CC produce antibodies. The polypeptides can also be used in vaccine
CC formulations, and to identify agonists and antagonists. The polypeptides,
CC antibodies, agonists and antagonists (which are bacteriostatic) are used
CC for the treatment and prevention of diseases such as otitis media in
CC infants and children, pneumonia in elderly, sinusitis, nosocomial
CC infections and invasive diseases, and chronic otitis media with hearing
CC loss. The polypeptides, agonists and antagonists are also used for
CC screening of antibacterial drugs. The BASB034 products of the invention
CC can be used screen for new antibacterial compounds that may target
CC resistant bacteria.

XX Sequence 442 AA:

Query Match 99.6%; Score 2354; DB 21; Length 442;
Best Local Similarity 99.5%; Pred. No. 6.8e-218;
Matches 440; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKVSLSTLSTLPCFAITAIQQAQAVPVAFVDEVSENDLGDNDLPIDVOSATOSA 60
DB 1 mkvslstlslstlpcfallalqgaqavpvaafvdevsendlgdndelpdvgsatgsa 60
QY 61 STDTPNPIDEHEPELYTTALENKMILNCSANODIMRLACDYLTVHGETPAVITKRSI 120
DB 61 stdtlnpidehepeilyttalenkmlncsainodimrlacdyltvhgetpaviktrsi 120
QY 121 RLDETIWOTIKGKPOVVOYQETTDPIFLMGNEKGMLTRKDAKOLEYAAKOFPLSLSPDLD 180
DB 121 rldetiwtikgkpvovyoettdpiflmgnekgmtlrrkdakoleyaakofplslsfdld 180

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:17:00 ; Search time 72.04 Seconds

(without alignments)
454,475 Million cell updates/sec

Title: US-09-787-083-8

Perfect score: 2363
Sequence: 1 MKVSLSTLTLSILPCFAILA.....YNHEATSPGVGLMNDMMGL 442

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /SIDSB/gcgdata/geneseq/AA1980.DAT:*
2: /SIDSB/gcgdata/geneseq/AA1981.DAT:*
3: /SIDSB/gcgdata/geneseq/AA1982.DAT:*
4: /SIDSB/gcgdata/geneseq/AA1983.DAT:*
5: /SIDSB/gcgdata/geneseq/AA1984.DAT:*
6: /SIDSB/gcgdata/geneseq/AA1985.DAT:*
7: /SIDSB/gcgdata/geneseq/AA1986.DAT:*
8: /SIDSB/gcgdata/geneseq/AA1987.DAT:*
9: /SIDSB/gcgdata/geneseq/AA1988.DAT:*
10: /SIDSB/gcgdata/geneseq/AA1989.DAT:*
11: /SIDSB/gcgdata/geneseq/AA1990.DAT:*
12: /SIDSB/gcgdata/geneseq/AA1991.DAT:*
13: /SIDSB/gcgdata/geneseq/AA1992.DAT:*
14: /SIDSB/gcgdata/geneseq/AA1993.DAT:*
15: /SIDSB/gcgdata/geneseq/AA1994.DAT:*
16: /SIDSB/gcgdata/geneseq/AA1995.DAT:*
17: /SIDSB/gcgdata/geneseq/AA1996.DAT:*
18: /SIDSB/gcgdata/geneseq/AA1997.DAT:*
19: /SIDSB/gcgdata/geneseq/AA1998.DAT:*
20: /SIDSB/gcgdata/geneseq/AA1999.DAT:*
21: /SIDSB/gcgdata/geneseq/AA2000.DAT:*
22: /SIDSB/gcgdata/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2363	100.0	442	21	AAV85271 BASB034 amino acid
2	2354	99.6	442	21	AAV85268 BASB034 amino acid
3	2348	99.4	442	21	AAV85270 BASB034 amino acid
4	2339	99.0	442	21	AAV85269 BASB034 amino acid
5	822	34.8	370	21	AAV75156 Neisseria meningit
6	815	34.5	370	21	AAV75157 Neisseria meningit
7	815	34.5	370	21	AAV75158 Neisseria meningit
8	815	34.5	374	21	AAV70629 Neisseria meningit
9	812	34.4	375	21	AAV70628 Neisseria meningit
10	246.5	10.4	355	19	AAV98871 H. pylori GHP 172
11	239.5	10.1	356	19	AAV10960 H. pylori ORF 07ap

12	156.5	6.6	253	18	AAV20760 H. pylori Outer me
13	117	5.0	1686	19	AAV70991 Human class II p13
14	108.5	4.6	800	8	AAV70420 Sequence encoded b
15	108.5	4.6	822	13	AAV26021 Alkaline cellulase
16	107.5	4.5	824	18	AAV23180 Bacillus sp. KSM-S
17	107	4.5	157	18	AAV20358 H. pylori Outer me
18	106	4.5	798	21	AAV40925 Human ORF1/ORF689
19	106	4.5	798	21	AAV90225 Human PGC-1 protei
20	106	4.5	798	22	AAV83944 Peroxisome prolif
21	102.5	4.3	537	22	AAV91000 C glutamicum prote
22	100.5	4.3	1726	18	AAV38756 Phosphatidyl inos
23	98.5	4.2	467	20	AAV13378 Amino acid sequenc
24	98.5	4.2	467	22	AAV94860 Human protein clon
25	98.5	4.2	467	22	AAV88401 Human membrane or
26	98.5	4.2	467	22	AAV80246 Human PRO260 prote
27	95.5	4.0	502	22	AAV90031 C glutamicum prote
28	95.5	4.0	516	22	AAV78876 C. glutamicum SRT
29	95.5	4.0	682	17	AAV04359 S. pneumoniae peni
30	95	4.0	522	22	AAV95616 Human protein sequ
31	95	4.0	888	22	AAV70751 S cerevisiae apopt
32	94	4.0	372	16	AAV69607 Gonococcal porin-5
33	94	4.0	761	20	AAV99084 Non-B, non-C, non-
34	94	4.0	1227	22	AAV81501 S. epidermidis ope
35	94	4.0	1717	22	AAV20498 Ostrinia nuda115
36	93.5	4.0	467	19	AAV75057 Human secreted pro
37	93	3.9	600	22	AAV75092 Human colon cancer
38	92.5	3.9	857	13	AAV29814 Cytohesin p1. My
39	91	3.9	1627	16	AAV67538 Human protein sequ
40	90	3.8	522	22	AAV95513 Modified dipteran
41	90	3.8	560	11	AAV06815 T. matsutake pyran
42	90	3.8	564	21	AAV10457 Trichoderma deriva
43	90	3.8	564	21	AAV81952 Human protein sequ
44	90	3.8	790	22	AAV93045 Mouse dephosphoryl
45	90	3.8	841	20	AAV24318

ALIGNMENTS

RESULT 1	
ID	AAV85271 standard; Protein: 442 AA.
XX	
AC	AAV85271:
XX	
DT	29-JUN-2000 (first entry)
XX	
DE	BASB034 amino acid sequence #4.
XX	
KW	Moraxella catarrhalis infection; diagnosis; staging;
KW	vaccine; bacteriostatic; treatment; prevention; pneumonia;
KW	sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW	hearing loss; antibacterial drug.
OS	Moraxella catarrhalis.
XX	
PN	MO200015802-A1.
XX	
PD	23-MAR-2000.
XX	
PF	14-SEP-1999; 99WC-EP06781.
XX	
PR	14-SEP-1998; 98GB-0020002.
XX	
PA	(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX	
PI	Ruelle J;
XX	
DR	WPI: 2000-271440/23.
XX	
PT	N-PSDB; AAA10703.
XX	
PT	Novel BASB034 polynucleotides and polypeptides from Moraxella
	catarrhalis used to prepare vaccines against bacterial infections

; Sequence 2, Application US/09355160D

APPLICANT: Malvar, Thomas M.
APPLICANT: Slater, Steven C.
APPLICANT: Spirdonov, Sergei
TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
FILE REFERENCE: 38-21(51847)B
CURRENT APPLICATION NUMBER: US/09/897,516
CURRENT FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/215, 161
PRIOR FILING DATE: 2000-06-30
NUMBER OF SEQ ID NOS: 8409
SEQ ID NO 4692
LENGTH: 2384
TYPE: PRF
ORGANISM: Xenorhabdus sp.
US-09-897-516-4692

Query Match 4.3%; Score 101; DB 5; Length 2384;
Best Local Similarity 20.8%; Pred. No. 0.93;
Matches 97; Conservative 58; Mismatches 124; Indels 188; Gaps 29;

QY 30 PVAFVDEVRN-----ENDLGODNDELPIVQSATQSASTDTANPLDEHEPELYTTA 79
DB 1015 PENFIDPTFRLKRTKAFNDLENDLGO----- 1040
QY 80 LENKTMILNCSALNODIMRLACYDTLVHGFTPAVITKRSIRLDETITWOTIKGKPOVYQ 139
DB 1041 -----SRLNMDLQFAIRLYLTERE-----RISNL--QLISG-----YI 1072
QY 140 ETTDPFILNGNEKMLTKDAKOLEYAAKQFTPLSLFDDLRNN--TPL-WS-----S 189
DB 1073 DGDGF--TSDRYHFIQKNNSEPEV--YWRITLDISL-RDSNNIISPLAMSEKTTLS 1125
QY 190 RPHNPVYLPIENHGRPN---RSPPTPSHEAQFTPNEFRABELKFOVSVKYKAADLW 245
DB 1126 LSGTLAIPIPIVYSGROYAIWVEREATPLMGADQ-KPTDYRAINVKF---TFKOSNGEW 1180
QY 246 GTDSDLM--FGYQOSHMOIFNGKNRPRFVHD-----YQPIFILQPVY--SDLP 292
DB 1181 SAPNELRIIDGDDANCEYPTKDGK---RIPDNEHYLKDEKYPKRALIMVVDQREEDP 1235
QY 293 WDQKVMIGMAVHNSGESAKLSRSM--NRAYIMAGMEKNTLVMPR-----I 339
DB 1236 W-----MGVLLYDTKE--KDSSTWVKNKQY---LELRDLLVDKTLIDDERKIVET 1282
QY 340 WGRIFKESSGSDPDNDPILDIYGYGDVRF--LYOLENSN-----ISQTVR 384
DB 1283 WYKLF-----NNPDLQHHYAGTEKEFVYKIEITKKNHFNFLKPPRSPILTLIDVR 1333
QY 385 YNPRSGKALQLDYVYPLGKIGISGYFOIFOGYOSLI--DYNHEATS 429
DB 1334 LDFTKTKLL-----MGR-----NSLIKFDNTNOAAS 1360

RESULT 5
US-09-801-368-52
; Sequence 52, Application US/09801368
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: Norman, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147

CURRENT APPLICATION NUMBER: US/09/801,368
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 09/487,558
PRIOR FILING DATE: 2000-01-19
PRIOR APPLICATION NUMBER: US 60/160,587
PRIOR FILING DATE: 1999-10-20
NUMBER OF SEQ ID NOS: 440
SOFTWARE: PatentIn version 3.0
SEQ ID NO 52
LENGTH: 1478
TYPE: PRF
ORGANISM: Saccharomyces cerevisiae
US-09-801-368-52

Query Match 4.1%; Score 97.5; DB 5; Length 1478;
Best Local Similarity 23.0%; Pred. No. 0.97;
Matches 110; Conservative 50; Mismatches 175; Indels 143; Gaps 26;

QY 2 KVISITLIS-----ILSCFAILAIQAKAVPNPVAIVDEVRSENDLGODNELPI---DV 53
DB 912 KVNRSNSTVSTNSIFYSPSPLKRGSKRYVSTSAAD-IFENDITFADAPPPDSDD 970
QY 54 QSATQSASTDTANPLDEHEPELYTTALENK--TMLINCSALNODIMRLACYDTLVHGFTP 111
DB 971 SDDSSSSDOLWKKKTAPR---TNNEKKDEKSDNSHTSDEIF---YDS---QTQ 1019
QY 112 AVITKRSIRLDETITWOTIKGKPOVYQETDPIFLNGNEKMLTKDAKOLEYAAKQFT 171
DB 1020 DKMERK-----MTFRSPPEVYQ-NLEKFFPRANLDKPYT-----EGIASPTS 1061
QY 172 PLISFPLDRNT-----PLMSSRPINPMVYLPIFNHG-----KPNRSPTP----- 213
DB 1062 PKSLDLSLPKNAVASSRTEBTPSRPVPPDSSYEFITDGLGKPKPLNQAATPRRTTIR 1121
QY 214 --SHEAQFTPNEFRABELKFOVSVKYKAEE-DLMDGTDSDLMFGYTOOSHMOIFNGKNSR 270
DB 1122 TIMEBASLARKN-----SVKLKRONTKMKGT--RMVEVTENHMYISIKAKNSK 1167
QY 271 PFRVHDYQPELFILOPYSDLPWDGKVMIGMA-----VHNSGESAKLSR----- 317
DB 1168 -----GEYKEFAW--MKGEMIGKGSFGAVYLLCLNTVTGEMMAVVKQVEVPKY 1211
QY 318 -SWNRATLMAGMEKNTLVMPRIRNGRIKFESSGSDPDNDPILDIYGYGDVRFILQLENK 376
DB 1212 SSOEATLIS-----TV-----EALRSEVSTLKLDHLNITVOYLGF-----ENK 1249
QY 377 SNISGTVRYNPRSGKALQLDYVYPLGKIGISGYFOIFOGYOSLIDYNHEATSEFVG 434
DB 1250 NNIT-----SLFLEYV--AGGSVGLIRMTGRDEPILI--KHLTYVLKGL 1291

RESULT 6
US-09-487-558-52
; Sequence 52, Application US/09487558
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: Norman, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/487,558

;; CURRENT FILING DATE: 2000-01-19
;; PRIOR APPLICATION NUMBER: US/09/801,368
;; PRIOR FILING DATE: 2001-03-07
;; PRIOR APPLICATION NUMBER: US 09/487,558
;; PRIOR FILING DATE: 2000-01-19
;; PRIOR APPLICATION NUMBER: US 60/160,587
;; PRIOR FILING DATE: 1999-10-20
;; NUMBER OF SEQ ID NOS: 440
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 52
;; LENGTH: 1478
;; TYPE: prt
;; ORGANISM: Saccharomyces cerevisiae
US-09-487-558-52

Query Match 4.1%; Score 97.5; DB 5; Length 1478;
Best Local Similarity 23.0%; Pred. No. 0.97;
Matches 110; Conservative 50; Mismatches 175; Indels 143; Gaps 26;

QY 2 KVSLSLTLSS-----ILSCFALIAIOAKAVPNPVAFVDEVRSENDLGODNELPI---DV 53
DB 912 KVRNSRSTVSTNSIFYSPLKRGNSKRVVSTSAAD-IFENDITRADAPMDSDD 970
QY 54 OSATOSASTDTANPLDEHEPELYTALLENK--TMLNSALNODIMRLACYDTLVHGTP 111
DB 971 SDDSSSSDDIIMSKKTAPE---TNNENKDKESDSSSTHSDEIR---YDS---QYQ 1019
QY 112 AVIKTKRSIRLEDTIMQTKGKPOVYVOETDPIFLMGNEKGLTKKAKOLEYAKOFT 171
DB 1020 DKMERK-----MTREPSREVOYQ-NLEKFFRAULDKPT-----EGIASPTS 1061
QY 172 PLTSLFDDLRNNT-----PLWSSRPHNPMYVLPIMHG-----KPNRSPNP----- 213
DB 1062 PSLDLSLSPKVVASRSTPSPRPVPPDSSEFIQDOLNGKNKRLNDAKPTKRTKTR 1121
QY 214 --SHEARQTPNEFRAPBELKFOVSAYKAAE-DLMGTDSDLMFGYTOQSHWOIFGKNSR 270
DB 1122 TTAHEASLARKN-----SVKLKRONTKMGT---RMVEVENHMHVSINKAKNSK 1167
QY 271 PRVVDYQEIFLQPVYSDLPMDKGVRMIGCA-----VHNSGESAKLSR----- 317
DB 1168 -----GEYKEFAM-MKGEIMIGSGFAGVYVLCVLWTTGEMAAKQVEVPKY 1211
QY 318 -SMNRAYLMAGNEMKMLTYMPRIWGRIFKEGSGSQPDNDPILIDYGYGQVRELYOLENK 376
DB 1212 SSONEALIS-----TV-----EALRSEVSTLKDLHLNITVQYLG-----ENK 1249
QY 377 SWISGTVARNPNSGKALOLDVYVPLGKISGTFQIFQGYGSLIDYHNEATISFGVGL 434
DB 1250 NNIY-----SLPLEYV--AGGSVGLIRMYGRFDEPLI--KHLTTOVLKGL 1291

RESULT 7
US-09-463-488A-63
;; Sequence 63, Application US/09463488A
;; GENERAL INFORMATION:
;; APPLICANT: Okamoto, Hiroaki
;; APPLICANT: Nishizawa, Tsutomu
;; TITLE OF INVENTION: Non-B, Non-C, Non-G Hepatitis Virus Gene, Polynucleotide, Polypeptide, Virus Particle, Method for Isolating Virus Particle, and Method
;; FILE REFERENCE: 45455-239360 20111-0040
;; CURRENT APPLICATION NUMBER: US/09/463,488A
;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: PCT/J998/03340
;; PRIOR FILING DATE: 1998-07-27
;; PRIOR APPLICATION NUMBER: JP 10-82962
;; PRIOR FILING DATE: 1998-03-13
;; PRIOR APPLICATION NUMBER: JP 9-314196
;; PRIOR FILING DATE: 1997-10-09
;; PRIOR APPLICATION NUMBER: JP 9-233246
;; PRIOR FILING DATE: 1997-07-25

;; NUMBER OF SEQ ID NOS: 64
;; SEQ ID NO 63
;; LENGTH: 761
;; TYPE: prt
;; ORGANISM: non-B, non-C, non-G hepatitis virus
US-09-463-488A-63

Query Match 4.0%; Score 94; DB 5; Length 761;
Best Local Similarity 20.7%; Pred. No. 0.77;
Matches 98; Conservative 62; Mismatches 178; Indels 136; Gaps 27;

QY 30 PVAFVDEVRSENDLGODNELPIDVOSATOS-----ASTDTANPLDEHEPELYTALLENK 83
DB 229 PKLFEDKWTYQDL---CKVPLVITATADLRYPSPCSQPTNNPC-----TFQVLRKN 280
QY 84 TMLNSALNODI---MRLACYDTLVH-----GEPNAYIK--TKRSIADDTIQ 128
DB 281 NTVIGTSVAKDQESTQDFEMVLYKTDSHYQTFATEAOLGRIPAPNPDGKTNT---KOOSWQ 337
QY 129 ---TIKGPQV---VYOETDPIFLMGNEK--GMLTKKAKOLEYAAKQFTPLSLFPL 179
DB 338 DNMSKNSPMTGNSGTYPOTTSEMYKIPYDSNNGFPTIYAKOXYLLERQ-----CNFN 392
QY 180 DRNNTPLMSSRPHNPMYVLPIMHGKPNRSPNTPSHARQF-----TPNEFRAP 229
DB 393 EVNN-----FVSKKVPQPSPTTPTVYDEYEHGWFENIFIGNRY---N 434
QY 230 LKQVVS-----YKVAEDLM-----CTDSD-----LM---FGYTOQ 258
DB 435 LQFOTAYVDITYPPLMDKGNKIMFOYLSKTKTDNEKOCYCTLEDMLALCFEYTDY 494
QY 259 SHQIENGKNSRFRVNDYQEIFLQPVYSDLPMDKGVRMIGCAVHNSGESAKLSRS 318
DB 495 VERQL--GRVVD---HEHAGLIMICP-YTOPRMDKRNPMNGYVYOTNFGNGAMPSC 547
QY 319 WNRAYLMAGNEMKMLTYMPRIW-----GHIFEGSGSQPD--NDPILIDYGY-----G 365
DB 548 SCQVPYVWQCRWR-----PALMFQOQOVLNDSITGPAYVDEKYNQVLTLYYNEIFNMWG 602
QY 366 DVFELVQLEKSNISGTVARNPNSGKALOLDVYVPLGNG---ISGYQIFQGY 416
DB 603 DMYPVOVKNPCGDSGIV---PQSGRPTREVQVSPLSMKPAYIFHFDSRRGF 653

RESULT 8
US-09-897-516-7203
;; Sequence 7203, Application US/09897516
;; GENERAL INFORMATION:
;; APPLICANT: Corbin, David R.
;; APPLICANT: Goldman, Barry S.
;; APPLICANT: Hunkle, Gregory J.
;; APPLICANT: Huesing, Joseph E.
;; APPLICANT: Krasomil-Osterfeld, Karina C.
;; APPLICANT: Malvar, Thomas M.
;; APPLICANT: Slater, Steven C.
;; APPLICANT: Spiridonov, Sergei
;; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
;; FILE REFERENCE: 38-21(51847)B
;; CURRENT APPLICATION NUMBER: US/09/897,516
;; PRIOR FILING DATE: 2001-06-29
;; PRIOR APPLICATION NUMBER: US 60/215, 161
;; PRIOR FILING DATE: 2000-06-30
;; NUMBER OF SEQ ID NOS: 8409
;; SEQ ID NO 7203
;; LENGTH: 2590
;; TYPE: prt
;; ORGANISM: Xenorhabdus sp.
US-09-897-516-7203

Query Match 3.9%; Score 91; DB 5; Length 2590;
Best Local Similarity 20.9%; Pred. No. 8.7;

Matches 68; Conservative 48; Mismatches 88; Indels 122; Gaps 18;

QY 1 MKVSLTSLTSLSCFAALLAIOAKAVNPVAVDEVR-----SENDL-----G 44
DB 156 LKNGSTMLSELPQY-LIRVKKSRYYLNFNVEDKNYLNFLKKEITLQOLPPYYHNGK 214
QY 45 ODNE---LPIDVOSATOSASTDTPANPLDEHEPELYTTALENKTMILNCSALNODIMRLAC 101
DB 215 SDNQTLTLPESVKN-----DLRFECHEKNKNL-----242
QY 102 YDTLVGEPYPAVKTKRSIRLDETIMQTKGKQVYVYQETTDPIFLMGNEKMLTKKDAK 161
DB 243 -----AVFSTLSTLSTMAE-----LTGQOELVFNNTT-----HGRTTKSEKY 278
QY 162 QLETAAROF---TPLSFPLDNRNTPMLSSRHPNPMYVLPFPMHCK-PNRSPTPSHEA 217
DB 279 VVGOMANTPVHCPISRTSSIIKOIKOVESALKNS-----YHGHGFPN-----SLT 325
QY 218 RQTPNEFRAPEL---KFQVSVKYKAAEDLMGTDSLMFGYTOOSHOIIFNGKNSRP- 271
DB 326 RLANGOGFSLPSILILYDRSAS-----DLETD-----QIOHFIHIDEFININPI 369
QY 272 -FRV--HDYQPEIFLT---QPVYSD 290
DB 370 LFRKLDHGYDOELKLTIOYLQTYFSD 395

RESULT 9

US-09-815-242-4989
; Sequence 4989, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4989
; LENGTH: 742
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-4989

Query Match 3.8%; Score 88.5; DB 5; Length 742;
Best Local Similarity 24.2%; Pred. No. 2.4; Indels 93; Gaps 23;
Matches 86; Conservative 41; Mismatches 136; Indels 93; Gaps 23;
QY 47 NELPIDVOSATOSAS-----TDTANPLDEHEPELYTTALENKTMILNCSALNO---DIMR 98
DB 187 NEHPARLYPNQGFASHFPGYTKKANPDDKKG-----LVGAMGLEQTYNDILS 234
QY 99 ---LACYDVLVHGET-PAVIKTKRSIRLDETIMQTKGKQVYVYQETTDPIFLMGNEK- 152
DB 187 NEHPARLYPNQGFASHFPGYTKKANPDDKKG-----LVGAMGLEQTYNDILS 234

DB 187 NEHPARLYPNQGFASHFPGYTKKANPDDKKG-----LVGAMGLEQTYNDILS 234
QY 99 ---LACYDVLVHGET-PAVIKTKRSIRLDETIMQTKGKQVYVYQETTDPIFLMGNEK- 152
DB 235 GTDGRVTFEKDIYNALGTAEEKKAVDGDITTLDSR---LONTLEDLMTQVENEKY 290
QY 153 -----GMLTKKDAKOLEYAAKQFTPLSLSPD-----LDRNNTPLWSS---RPHNPMY 196
DB 291 EPVSMTAML--MEAKTGEIVAMSQR---TENPETKQGLDNGT--WQNLVESYEPGKS 343
QY 197 VLPFPMHGRKPSRNPSPSHARQFTPNEFRAPELKFQVSVKYKAAEDLMGTDSLMFGCYT 256
DB 344 TIKLF-----TTAASMEQGFNPMLFNRVGIGIYG-----DVTVDHD---YT 384
QY 257 QOSHWQIFNGKNSRPFV-VHDYQPEIFLT---QPVYSDLPMDGKYMIGMCAVHNS--NGE 311
DB 385 R-----LNGKEYLYNTRALISWSSNIGKVKLEQKMGDEKMEYIKKGFSTHSGLSGE 438
QY 312 SA-XL--SRSWNRVAYLAGMEKMLTV---MPRWGRIFEGSGSOPDDNPDLID 360
DB 439 SAGKLPQTNFVDRMBSAFG---QALITYNFMGMGFSALANDGSMLOPHYISKIYD 491

RESULT 10

US-09-815-242-10634
; Sequence 10634, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10634
; LENGTH: 742
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10634

Query Match 3.8%; Score 88.5; DB 5; Length 742;
Best Local Similarity 24.2%; Pred. No. 2.4;
Matches 86; Conservative 41; Mismatches 136; Indels 93; Gaps 23;
QY 47 NELPIDVOSATOSAS-----TDTANPLDEHEPELYTTALENKTMILNCSALNO---DIMR 98
DB 187 NEHPARLYPNQGFASHFPGYTKKANPDDKKG-----LVGAMGLEQTYNDILS 234
QY 99 ---LACYDVLVHGET-PAVIKTKRSIRLDETIMQTKGKQVYVYQETTDPIFLMGNEK- 152
DB 187 NEHPARLYPNQGFASHFPGYTKKANPDDKKG-----LVGAMGLEQTYNDILS 234

Db 235 GTDGRVYFEKDIYGNALPCTVAEKKAVDGDITTTLSR-----LQNTLEDLMTQVENEK 290
Qy 153 -----GMLTKKDAKOLEYAAKQFTPLSLSPD-----LDRNNTPLMSS-----RPHNPMY 196
Db 291 EYVSMATML-NEAKTGEIVAAVSQRP---TFNPETRKQGLDDNGT--NQGLLVESPEPEPS 343
Qy 197 VLPITMHGKPNRSPNTPPSHEARQFTPNNEFRAPELKFQVSVKKAADLMTGTSDDLFGYTT 256
Db 344 TKLFL-----TTAASMEGQFNPENLEFNNGVGIQVG-----DVTVDNHD-----YT 384
Qy 257 QOSHMOIFNGKNSRPER-VHDYQPEIFELT--QPVYSDLPMDGKVRMIGAVHHS--NGE 311
Db 365 R-----LNGKELYLYRQAISSNIGAVKLEQKMGDEKMEYLFKFGFTSTHSGLSGE 438
Qy 312 SA-KL--SRSMNRAYLMAGMEKNLTV---MPRIWGRIFKEGSGSQPDNDPDIID 360
Db 439 SAGKLPCTNFVDRAMSAGF---QAITVTNFMQMKFSALANGSMLOPHYISKIYD 491

RESULT 11
US-09-897-516-7356
Sequence 7356, Application US/09897516

GENERAL INFORMATION:

APPLICANT: Corbin, David R.
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Huesling, Joseph E.
APPLICANT: Krasomil-Osterfeld, Karina C.
APPLICANT: Malvar, Thomas M.
APPLICANT: Slater, Steven C.
APPLICANT: Spitzdonov, Sergei
TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
FILE REFERENCE: 38-21(51847)B
CURRENT APPLICATION NUMBER: US/09/897,516
CURRENT FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/215, 161
PRIOR FILING DATE: 2000-06-30
NUMBER OF SEQ ID NOS: 8409
SEQ ID NO 7356
LENGTH: 678
TYPE: PRF
ORGANISM: Xenorhabdus sp.
US-09-897-516-7356

Query Match 3.7%; Score 86.5; DB 5; Length 678;
Best Local Similarity 20.0%; Pred. No. 3.2;
Matches 56; Conservative 32; Mismatches 85; Indels 107; Gaps 13;

Qy 128 QTIKGRPOVVOETT--DPIF--LMGNEKML-----TKKDAKOLEYAAKQFTPL 173
Db 62 KTIPIFSQVRSCTGNDPVRFGMFGSLRLMLDDGEVMGSCSRMDAPSSYISPESDYVL 121
Qy 174 SLSPFLDRNNTPLMSSRPNRPVLPPIFMHGKPNRSPNTPPSHEARQFTPNNEFRAPELKFQ 223
Db 122 SL--VKGPQTVLMG-----PNSAGTIRERSR---PQSEPAKGN 158
Qy 234 VSVKKAADLMTGTSDDLFGYTOOSHMOIFNGKNSRPFVHDYQPEIFLTQPVYSDLPW 293
Db 159 IST-VAGSGNRDGNADLSFG--NQYGMRLIGKNS---RSNDYK----- 197
Qy 294 DQKVRMIGAVHHSNGESAKLSRSNMRAYLMAGMEKNLTVMPRIWGRIFKEGSGSQPD 353
Db 198 -----DGDGVRIPSHMNK-----WNTDVAL-----GWTDP 222
Qy 354 DNPDLIDYVYGDVRFVLEKNKSNISGTVRYNPRSGKA 393
Db 223 DNTLELNLNGGD-----GKARYAGRAMDS 248

RESULT 12
PCT-US01-27760-801
Sequence 801, Application PC/TUS0127760

GENERAL INFORMATION:
APPLICANT: Hysq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-115/ 785
CURRENT APPLICATION NUMBER: PCT/US01/27760
CURRENT FILING DATE: 2001-10-11
PRIOR APPLICATION NUMBER: 09/687,527
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 892
SOFTWARE: Custom
SEQ ID NO 801
LENGTH: 1021
TYPE: PRF
ORGANISM: Homo sapiens
PCT-US01-27760-801

Query Match 3.6%; Score 84.5; DB 1; Length 1021;
Best Local Similarity 19.7%; Pred. No. 8.8;
Matches 68; Conservative 44; Mismatches 100; Indels 133; Gaps 17;

Qy 131 KGRPOVVOETTDPPIFLMGNEKMLTKKDAKOLEYAAKQFTPLSLSPDLDNRNTPMW-SS 189
Db 499 KAKAQVYLQRLD-----VQEIFHKROYSLMKLAKQTRPVQ--PVAPRESSPKVSS 550
Qy 190 RPHNPMYVLPPIFMHGKPNRSPNTPPSHEARQFTPNNEFRA-----PELKFQVSVKKAEDL 244
Db 551 KTSQPSSTVPL--ARPLRTSEEP-----YETELNSRCKEDEDKFEVK-----SEEL 596
Qy 245 WGTSDLMFGYTOOSHMOIFNGKNSRPFVHDYQ-----EFLFQPVYSDLPMDGKVR 298
Db 597 F-----ESHHERNPPELEQOARGLDLSPPRRIRIIDLLEETL----- 634
Qy 299 MIGAVHHSNGESAKLSRSNMRAYLMAGMEKNLTVMPRIWGRIFKEGSGSQPDNDPDI 358
Db 635 -----IKEIKSIIDGY-----ITPMEDIW---LKHILPDLQNNKDF 668
Qy 359 LDYVYGDVRFVLEKNK-----KSNISGTVRYN---PRS-- 389
Db 669 L-----FGNIRELYEFHNRPTLKELEKCAENPELLAHCFLEKRDQIYFYKHRLPRARA 724
Qy 390 -----GKGAQLDQVYVPLGKIGISYRQIFQYQGSGLDY 423
Db 725 IMQECDCAYEGVCORQLDHLPL-----FYLKGPSQRLIKY 762

RESULT 13
US-09-897-516-6909
Sequence 6909, Application US/09897516

GENERAL INFORMATION:

APPLICANT: Corbin, David R.
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Huesling, Joseph E.
APPLICANT: Krasomil-Osterfeld, Karina C.
APPLICANT: Malvar, Thomas M.
APPLICANT: Slater, Steven C.
APPLICANT: Spitzdonov, Sergei
TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
FILE REFERENCE: 38-21(51847)B
CURRENT APPLICATION NUMBER: US/09/897,516
CURRENT FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/215, 161
PRIOR FILING DATE: 2000-06-30
NUMBER OF SEQ ID NOS: 8409
SEQ ID NO 6909
LENGTH: 802
TYPE: PRF
ORGANISM: Xenorhabdus sp.
US-09-897-516-6909

Query Match 3.5%; Score 83.5; DB 5; Length 802;

Best Local Similarity 19.0%; Pred. No. 7.7;
Matches 67; Conservative 59; Mismatches 109; Indels 117; Gaps 20;

```
QY 147 LMGNEKMLTKKADKOLEYAKOFTPLSLFDLRNTPMLSSRPHNMYLPIFMHGRP 206
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 41 LVGLEKALTSSENPVVKAEK---MGKLFDDIKINNNTLT---EDIALANFL----- 89
QY 207 NR-----SPNTSHEAROTPNDEFRAPELKFQVSVKKAEDLMGSDLSMGSTQOSH 260
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 90 TRLLCFEADTGIIFSDAQT---AIKSYTEED--GSDLLAFL----- 128
QY 261 WOJFNKNSRPFVHQPEIFLTPQPYSDLPMDGKVRMIGMAVHNSGASAKLSRWN 320
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 129 YQFSVLNT---PESDTRINLPKRI-----NNFRYN-GDLSSDDPIPELKKGR 176
QY 321 RAYLMAG-MEMKNLYVPRIMGRIFKSGSGSDPDNDILD-----YYGQGVRF 369
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 177 RILLDCGSMDSSEIN--POIFGSMFOA-----VIDEGRSLGQHYTSNIMK 223
QY 370 LYQ-----LEKNSISGTVR-----YNPSSGKG-----ALQLD 397
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 224 VIQPLEFLEPLRSDPEKRRKRVNGLAKLVRIKIKYFDPACGSGCNFLIVAYKELRLME 283
QY 398 YVYPLGK-----GISGYFOIFQGYGOSLIDYHNHEATSFVGL-----MLNDM 439
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 284 IIMALGEIDPQSIPLMSG-LHLSQFYGIETDPFACFVARLSIMLAHQJLNNQW 334
```

RESULT 14

US-09-897-516-5307

; Sequence 5307; Application US/09897516

; GENERAL INFORMATION:

; APPLICANT: Corbin, David R.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Huesling, Joseph E.

; APPLICANT: Krasomil-Osterfeld, Karina C.

; APPLICANT: Malvar, Thomas M.

; APPLICANT: Slater, Steven C.

; APPLICANT: Spilidonov, Sergei

; FILE REFERENCE: 38-21(51847)B

; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof

; CURRENT APPLICATION NUMBER: US/09/897,516

; PRIOR FILING DATE: 2001-06-30

; PRIOR APPLICATION NUMBER: US 60/215, 161

; NUMBER OF SEQ ID NOS: 8409

; SEQ ID NO 5307

; LENGTH: 595

; TYPE: PRT

; ORGANISM: Xenorhabdus sp.

US-09-897-516-5307

Query Match 3.5%; Score 82.5; DB 5; Length 595;

Best Local Similarity 23.4%; Pred. No. 6.1;

Matches 50; Conservative 29; Mismatches 76; Indels 59; Gaps 13;

```
QY 30 PVAFEVRESENDLGONDELPIDVOSA---TOSASTDTANPLDHEPELYTTALENKT 85
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 186 PSSAVNTSTVDTP-----SVDTNSVDSITTAIVETDQANTADTTETDM-----TGQF 232
QY 86 LINCASLADIMRLACYDT-----LVHGETPAVTKTRISIRLDETIMQITNG---KP 134
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 233 TINSVPVNHQVVKI--YDTOSPLANLLAQAEQGV-----TLVVGPLKKP 276
QY 135 QV--VQETIDPFLMGNEKMLTKKADKOLEYAKOFTPLSLFDLRNTP--LWSSR 190
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 277 QVEQLVQINT-PLNTLA---LNKPTTQLH---PNFCYFSLSPDEAKSMAQIWOQ 327
QY 191 PHNPMYVLPFIEMHGKPRSPNTPSHEAROTPN 224
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 328 KHNPLVLVPRSVLG--SRVANAFATEWQJLGSGE 359
```

RESULT 15

US-09-800-198-72

; Sequence 72; Application US/09800198

; GENERAL INFORMATION:

; APPLICANT: Vernet, Cornie AM

; APPLICANT: Fernandes, Elma

; APPLICANT: Shimkels, Richard A

; APPLICANT: Hermann, John L

; APPLICANT: Majumder, Kunud

; APPLICANT: Mishra, Vishna

; APPLICANT: Mezes, Peter S

; APPLICANT: Rastelli, Luca

; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: 15966-697

; CURRENT APPLICATION NUMBER: US/09/800,198

; CURRENT FILING DATE: 2001-03-05

; PRIOR APPLICATION NUMBER: 60/186,596

; PRIOR FILING DATE: 2000-03-03

; NUMBER OF SEQ ID NOS: 98

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 72

; LENGTH: 2765

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-09-800-198-72

Query Match 3.5%; Score 82.5; DB 5; Length 2765;

Best Local Similarity 20.4%; Pred. No. 58;

Matches 99; Conservative 58; Mismatches 156; Indels 173; Gaps 27;

```
QY 57 TOSASTDTANPLDHE-----EPELYTTALENKTMLINCASLADIMDI 96
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1609 TYSADNDVTELIDNNNGSLKTRDSSGMPRLHLPMDNQIITLVGTNGGLAAVSTQNDL- 1667
QY 97 MRLACYDTLVHGETPAVTKTRISIRLDETIMQIT-----KKPQVYQET 141
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1668 LGLMTYD---GNT-GLIATKS---DEGTWTFYDYDHEGRLLNVTPTGVVTSIAREM 1718
QY 142 TDPIFL-----MGNEKMLTKKADKOLEYAKOFTPLSLFDLRNTPMLSSRPHNPM 195
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1719 EKSTIVDIENSNRDNDVTITNLSVEASTYVQ-----DOYRNSYQLCSNGLTLYM 1770
QY 196 YV--LPIFMHGKPRSPNTPSHEAROTPNDEFRAPELKFQVSVKKAEDLMGSDSLMF 253
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1771 YANGMGVSFHSPEHYLAGT-----LFTPTGRC-----NLSLPMEN-----GLNSIEMR 1813
QY 254 GYTOQSHWQIFNGKNS---RPRVH-----DYQPELFLTPQVYS-----LPMGK 296
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1814 LRKEQ-----IKGRVYTIQRLRVHGRMLLSIDYDRI-RTREKIYDDHRRKFTLRITLYDOV 1867
QY 297 VRMI-----GMGAHVHNS-----NGESAKLSRSMNNRAYLMAAGMKNLYVMPRIWGRIFK 347
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1868 GRPLMLPSSGLAAVWVSFFNGRLAGLQRC-----AMSERDIDKQGHVSRMPADG 1920
QY 348 SGSDPDNDPDLIDYGYGDV-----RFLYOLEKNSNIS----- 380
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1921 K-----VMSYSLYDKRMVLLLSQROYIFREYDSSDLHLAVTMPVSARHSMSTHS 1970
QY 381 -GTVR--YNPSSGALOLDY-----VYPLKGISGYQIIFQGYQO---SLIDYN 425
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1971 IGYIRNIYNPESNASVIFDYSDGRLIKTSFLGTG---RQVFPKYGKLSKLSLSEIYVDS 2026
QY 426 EATSPG 431
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 2027 TAVTFG 2032
```

Search completed: November 30, 2001, 14:25:08
Job time: 540 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:26:27 ; Search time 77.9 Seconds
(without alignments)
829.941 Million cell updates/sec

Title: US-09-787-083-6
Perfect score: 2360
Sequence: 1 MKVSLSTLTLSLSCFALTA.....YNHRTSRGCLMLNDMGL 442

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP:archaea:*
2: SP:bacteria:*
3: SP:fungi:*
4: SP:human:*
5: SP:invertebrate:*
6: SP:mammal:*
7: SP:mhc:*
8: SP:organelle:*
9: SP:phage:*
10: SP:plant:*
11: SP:rodent:*
12: SP:virus:*
13: SP:vertebrate:*
14: SP:unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	815	34.5	382	2 09K0U7	09K0U7 neisseria m
2	815	34.5	409	2 09J721	09J721 neisseria m
3	373.5	15.8	306	2 09C122	09C122 pasteurella
4	366	15.5	329	2 09PMU8	09PMU8 campylobact
5	360.5	15.3	292	2 09SID7	09SID7 yersinia ps
6	360	15.3	289	2 09L6N9	09L6N9 salmonella
7	356.5	15.1	292	2 09Z4N8	09Z4N8 enterobacte
8	342.5	14.5	297	2 03Z349	03Z349 campylobact
9	246.5	10.4	355	2 025241	025241 helicobacte
10	240.5	10.2	355	2 09Z1X5	09Z1X5 helicobacte
11	123	5.2	278	2 09XB53	09XB53 erwina car
12	114	4.8	1686	4 000443	000443 homo sapien
13	110.5	4.7	602	11 063485	063485 rattus norv
14	109.5	4.6	821	2 059241	059241 bacillus sp
15	108	4.6	824	2 09F216	09F216 bacillus sp
16	105.5	4.5	901	5 018749	018749 caenorhabdi
17	104	4.4	783	2 045554	045554 bacillus sp
18	104	4.4	798	4 09UN32	09UN32 homo sapien
19	104	4.4	798	4 09UBK2	09UBK2 homo sapien

20	103.5	4.4	3247	12 065553	065553 bovine herp
21	103	4.4	435	2 09X9C0	09X9C0 streptococc
22	101.5	4.3	403	5 091736	091736 hydra magni
23	101.5	4.3	1046	2 084941	084941 streptococc
24	101	4.3	4307	5 019319	019319 caenorhabdi
25	100.5	4.3	1509	11 061194	061194 mus musculi
26	100	4.2	719	2 057114	057114 streptococc
27	100	4.2	719	2 09RET8	09RET8 streptococc
28	99.5	4.2	719	5 09V9E6	09V9E6 drosophila
29	99.5	4.2	1658	11 061182	061182 mus musculi
30	99	4.2	719	2 054948	054948 streptococc
31	99	4.2	719	2 054946	054946 streptococc
32	99	4.2	719	2 09WVW0	09WVW0 streptococc
33	99	4.2	719	2 09WM11	09WM11 streptococc
34	99	4.2	719	2 09RET6	09RET6 streptococc
35	99	4.2	719	2 09RET4	09RET4 streptococc
36	99	4.2	727	2 09F2G4	09F2G4 streptococc
37	99	4.2	1372	10 09F192	09F192 arabidopsis
38	99	4.2	5005	2 09PP25	09PP25 ureaplasma
39	98.5	4.2	467	4 09UMT5	09UMT5 homo sapien
40	98.5	4.2	467	4 09BYT2	09BYT2 homo sapien
41	98.5	4.2	576	2 09L115	09L115 streptomyce
42	98.5	4.2	660	2 09K8K8	09K8K8 bacillus ha
43	98.5	4.2	788	10 09SCV5	09SCV5 arabidopsis
44	98	4.2	719	2 054947	054947 streptococc
45	98	4.2	719	2 054949	054949 streptococc

ALIGNMENTS

RESULT 1
ID 09K0U7 PRELIMINARY; PRT: 382 AA.
AC 09K0U7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PHOSPHOLIPASE A1, PUTATIVE.
GN NMB0464.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=2017575; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J., Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A., Mason T., Clecko A., Parksey D.S., Blair E., Clifton H., Clark E.B., Cotton M.D., Uitterlinden T.R., Khouri H., Qin H., Vamathevan J., Gill J., Scarlato V., Maignani V., Pizsa M., Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappelli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.";
RT Science 287:1809-1815(2000).
RL EMBL: AE002403; AAF40901.1; -;
DR TIGR: NMB0464; -;
DR InterPro: IPR003187; PLAL.
DR Pfam: PF02253; PLAL; 1.
KW Complete proteome.
SO SEQUENCE 382 AA; 42714 MW; B468A802F062E836 CRC64;

Query Match 34.5%; Score 815; DB 2; Length 382;
Best Local Similarity 44.7%; Pred. No. 1.6e-62;
Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;
OY 87 INCALNODIMRLACVDTLVHGERPAVY-----TKRSIRLDETITWOTI-KGKPOVYVOE 140
DB 33 LQCALNDVNTRLACVDRIFRAQLPSSAGDEGSKAVNLTEIVRSLSLKGAVIYVER 92

OY	141	TTDPFLMGNEKGMTLTKDKDALEFAAKQFPLTSFSPDLDRNN-TPLMSSRPHPNMYPLP	199
Db	93	GGDAL-----+-----PADSAGETADITPLSLMTLDRKDLGGLGVRHNPMYTLMP	139
OY	200	IFMHCKPNKSPNTPSHEAR-QETPNEFPAPELKFOVSVYKVAABEDLWGTSDLMWGYTQO	258
Db	140	LMYNSSPNVAPGSPRGTTVDEKFGQOKRAETKLGQVFSFKSLAEDLEFKTRADLMFGYTOR	199
OY	259	SHMWQFN-GKNSRPRFVNDYQPEIFLITQPVYDDLWMDCKVNRITGCAVNHSGESAKLSR	317
Db	200	SDMQLYNGRKAAPRNMTDYKEPFIETQPVKADLPFEGRLRLMAGFVHQSGSRPESR	259
OY	318	SWNRAYLMAAGMEKMLUTWMPRIWGRIFFKSGSSQDDNDLIDYGYGVPFLYOLENKS	377
Db	260	SWNRITMAAGMGKGLTYIPRYVVAFAFD-SSDK-NDNPDIADYMGYGVKIQYRLNDRQO	317
OY	378	NISGTVRYNPSRSGKALOLDVYVPLGKGISGYFOFGYQSGSLIDYNHEATSFVGGLMLN	437
Db	318	NYSVLRYNPKGYGAIEAAYTFPIKGLKGVRFHGESLIDYNHKONGIGIGLMPN	377
OY	438	DWNGL 442	
Db	378	DLDT 382	
RESULT	2		
ID	09J721	PRELIMINARY:	PRT: 409 AA.
AC	09J721.		
DT	01-OCT-2000 (TREMBLrel. 15, Created)		
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	PURATIVE PHOPHOLIPASE.		
GN	NMA2021.		
OS	Neisseria meningitidis (serogroup A)		
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.		
OX	NCBI_TaxID=65699;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=22491 / SEROGROUP A / SEROTYPE 4A;		
RC	MEDLINE=20222556; PubMed=10761919;		
RX	Path111 J., Achman M., James K.D., Bentley S.D., Churcher C., Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T., Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Jørgensen K., Leather S., Moule S., Mungall K., Quail M.A., Raftery D., Rutherford K.M., Skellern M., Skelton J., Whitehead M.A., Spratt B.G., Barrett B.G.		
RA	"Complete DNA sequence of a serogroup A strain of Neisseria		
RT	meningitidis 22491."		
RL	Nature 404:502-506(2000).		
DR	EMBL: AL162757; CAB85240.1; -		
DR	InterPro: IPR003187; Pfam: PF02253; Pfam: PF02253; Pfam: PF02253		
KW	Complete proteome.		
SO	SEQUENCE 409 AA; 45862 MW; CD6585B064D01A41 CRC64;		

```

Query Match Similarity      34.5%; Score 815; DB 2; Length 409;
Best Local Similarity      44.7%; Pred. No. 1.8e-62;
Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;

QY      87  INCSALNODIMELACYDYLVEHETPAVI-----KRRSRIRLDETTWQRI-KGKPOVYVOE 140
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      60  LQCALALDWDVTRFLACYCDRFRFAQNLPSRSMQGEQESKAVNLLETFRSSLDKREAVIYVER 119

QY      141 TTDPTFLMGNEKGMGLTKKDAKOLEYAAKQFTPLSLSFSDIDRNN-TPLMSSRPHNPYYLP 199
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      120 GGDAL-----PADSAGEFTADITYTPLSLMYDLDKNDLKGILGVRHEHNPYYLMP 166

QY      200 IIMHKKRPNRSPPTPSHEAR-QFTPNDFRFRPELKFQVSYVKAEDLDMGSDSLMGCTYQQ 258
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      167 LMYNNSPNTPAPSSPRGTTVOEKFQOQKAAERFKLQVSPFKSLAIEDLFKTRADLMGCTYQR 226

```

```

OY      259 SHMIOEF--GKNSRPRPHVDYOPEIFETLPQVSYDLDPMDGKVRAIMGAGVAHNSGCSEAKLSR   317
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      227 SMOQTLYNCRKAAPFRNRNDYKRPKEITLTQPVKADLFFGRRLRMGLGFVNHSNGCSRPESR   286
OY      318 SNMRAYILMAGMEMWKNLYVPRIWGRI FEKSGSSQPDPDNILLDYGYGDVRFYLLENKS    377
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      287 SMNRITYLAAMGWMLGTIVPVRVWAFAFDQ-SSDK-NNDNPEDIDAWMGYGCVKRKYQRYLRMDRQ   344
OY      378 NISGVRRNPBREGSKALODLVYPLPGKGKSSGYFEOFGCGSLTIDYNHEATSPFYGLMIN    437
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      345 NVYSVLRNPKTGTYGAIEAFTPIFKGLKGVYRGHFHGIESLLIDYHRNQNGSIGTLAFNN    404
OY      438 DWMGL 442
       |||
Db      405 DLEDGI 409
RESULT          3
OQC122         PRELIMINARY; PRT; 306 AA.
AC OQC122;
DT 01-JUN-2001 (TREMBLrel_17, Created)
DT 01-JUN-2001 (TREMBLrel_17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel_17, Last annotation update)
DE HYPOTHETICAL PROTEIN PM1426.
GN PM1426.
OS Pasteurella multocida.
OC Bacteria; Proteobacterla; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_Taxid=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida pm70." ;
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; AE006179; AAK03510.1; -.
DR InterPro: IPR003187; Pfam: PF02253; PLAI: 1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 306 AA; 35580 MW; 5AF3DBECIC22B26E CRC64;
```

Query Match	15.8%	Score 373.5	DB 2	Length 306
Best Local Similarity	37.6%	Pred. No.24e-24		
Matches	86	Conservative	43	Matches 91
				Indels 9
				Gaps 5

DE PHOSPHOLIPASE A (BC 3.1.1.32).
GN PLDA OR CUI351.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OX Campylobacter.
RN NCBI_TaxID=197;
RP SEQUENCE FROM N.A.
RC STRAIN=MCCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Kelley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
RA Jagers K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., Van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences."
RL Nature 403:665-668(2000).
RL EMBL: AL139078; CAB73778.1;
DR InterPro: IPR003187; Pfam:
DR Pfam: PF02253; Pfam: 1.
DR PRINTS: PR01486; PHPLIPASEA1.
DR Complete proteome.
SQ SEQUENCE 329 AA; 38880 MW; A32731F2B751AC44 CRC64;

Query Match 15.5%; Score 366; DB 2; Length 329;
Best Local Similarity 33.2%; Pred. No. 1.2e-23;
Matches 105; Conservative 48; Mismatches 125; Indels 38; Gaps 13;

OY 130 IKKRPVVYQETDPTFLMNEKMKLTKKAKOLEYAAKQFTPLSLFDDRRNT--PLM 187
DB 43 LKSSSVLISOEQNNSSQATOTONSITIKKEQ-----DFSRLALANYLGEMSEFNP 96
OY 188 SSRPNPMVYLPTFMHGKPRSPNTPSHARQFTPNFRAPELKFQYVYKAAEDLGT 247
DB 97 IS-SYKMNLYPL-FAYSFNGLGVNNKSEA-----KFOLSVKKRLFENLGL 141
OY 248 DSDLMFGYTOQSHWOIFNGKNSRPFRVHDYQPEIFLTQPVY-SDLPMGKVRMIGMAVH 306
DB 142 DEKYVYLAITQTSWMOIY--EHSFPFRETNYQPEFIDLPLYLKDYEFEFNMLR--VGLH 196
OY 307 HSKGESAK--LSRSNRAVYLMAGMEKMLVMPRIMGRIRKESGSDPDNDITLDYGY 364
DB 197 ESNKGKDEINQSSRMNRYVSTALYKFLFVPRLMYRI--PENKDDDPALHLMGN 253
OY 365 GDVRFYOLENKNISGTVRYNPR--SGKALQLDYVYPL-GKISGFOLFQGGSLI 421
DB 254 FDVNLAV-LGDDYFINMLMNNLKFHNKCAIQVDLGIENNGIYVYLDYFNCGESLI 312
OY 422 DYNHEATSEGVGLMLN 437
DB 313 DYNKHLQRLSTGFLIS 328

RESULT 5
OY9SID7 PRELIMINARY; PRT; 292 AA.
AC OY9SID7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE PHOSPHOLIPASE A.
GN PLDA.
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX Yersinia.
RN NCBI_TaxID=633;
RP SEQUENCE FROM N.A.
RC STRAIN=YPIII PIB1.
RA Karlyshev A.V., Winzeler E.A., Williams K.J., Oyston P.C.,
RA Tibball R.W., Wren B.W.;

RT "Biochip-based Signature-Tagged Mutagenesis: Identification and
RT Characterisation of Y. pseudotuberculosis gene plda essential for
RT virulence in mice."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ245393; CAB51586.1;
DR InterPro: IPR003187; Pfam:
DR Pfam: PF02253; Pfam: 1.
DR PRINTS: PR01486; PHPLIPASEA1.
SQ SEQUENCE 292 AA; 33758 MW; 8E712D908ACB6BA5 CRC64;

Query Match 15.3%; Score 360.5; DB 2; Length 292;
Best Local Similarity 39.7%; Pred. No. 3.e-23;
Matches 83; Conservative 34; Mismatches 86; Indels 9; Gaps 5;

OY 229 ELKFOVSVKYKAAEDLMDTSDLMFGYTOQSHWOIFNGKNSRPFRVHDYQPEIFLTQPV- 287
DB 86 EVKFOLSLAPPIWMGIDSLGASYSYORSWMOASSESSPRETINYEQLFLAWSTD 145
OY 288 YSDLPMDGKVRMIGMAVHNSGESAKLSRSNRAVYLMAGMEKMLVMPRIMGRIRKES 347
DB 146 YEIAGW--TFREVEFEFNHOSNGKADPTSRSWNRYVRYMAQRGNLEIDLKPMYRIEESD 203
OY 348 GSGQPDNDPILDYGYGDIYFQYOL-ENKSNISGTVRYNPRSGKALQLDYVYPLGKI 406
DB 204 S---KDNPDITTKMGYTRKLVGALGDSVPSLDG--RYNMNTSGGAEMGMSFPIKHY 258
OY 407 SGYFQIFQYGOGLIDYNHEATSEGVGLMLND 438
DB 259 RFTYQVFSGYGESMIDNFRQTRVGVGIMLND 290

RESULT 6
OY9L6N9 PRELIMINARY; PRT; 289 AA.
AC OY9L6N9;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE PLDA PROTEIN.
GN PLDA.
OS Salmonella typhimurium LT2.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX Salmonella.
RN NCBI_TaxID=99287;
RP SEQUENCE FROM N.A.
RC STRAIN=SGSC1412;
RA Washu;
RT "The Salmonella typhimurium Genome Sequencing Project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SGSC1412;
RA Waterston R.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF233324; AAF3435.1;
DR InterPro: IPR003187; Pfam:
DR Pfam: PF02253; Pfam: 1.
SQ SEQUENCE 289 AA; 32967 MW; DA97FE5E1651C49C6 CRC64;

Query Match 15.3%; Score 360; DB 2; Length 289;
Best Local Similarity 39.7%; Pred. No. 3.3e-23;
Matches 85; Conservative 29; Mismatches 90; Indels 10; Gaps 4;

OY 226 RAEPLKQVSVKYKAAEDLMDTSDLMFGYTOQSHWOIFNGKNSRPFRVHDYQPEIFLTQ 285
DB 83 RKDEVKFOLSLAPPIWMGIDSLGASYSYORSWMOASSESSPRETINYEQLFLG 142
OY 286 PV-YSDLPMDGKVRMIGMAVHNSGESAKLSRSNRAVYLMAGMEKMLVMPRIMGRIF 344
DB 143 ATDYRFAGW--TLRDVEMGYNHDSNGSDPTSRSMNRLYTRMAENGNMVLVEYKPMYVI- 199

RA Cotton D.D., Meldman J.M., Fujii C., Bowman C., Matthey L., Mallin E.
RA Hayes W.S., Borodovsky M., Kaip P.D., Smith H.O., Fraser C.M.,
RA Venter U.C.;
RT The complete genome sequence of the gastric pathogen *Helicobacter*
RT *pylori*.
RL Nature 388:539-547(1997).
DR EMBL: AE000564; AAD07564.1; -.
DR TIGR: HP0499; -.
DR InterPro: IPR003187; PLAI.
DR Pfam: PF02253; PLAI; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 355 AA; 42486 MW; 461960F1565AB0A CRC64;

Query Match	10.4%	Score 246.5;	DB 2;	Length 355;
Best Local Similarly	25.1%;	Pred. No. 3.2e-13;		
Matches 82; Conservative	41;	Mismatches 117;	Indels 87;	Gaps 11;

Qy	157	KKDAKOLEYAKORTPELSDDDDRNTPLMSSPHNPMVLPELMGKRNBSRPNTSHE	216
Db	69	KKYLNMDDLOCTYLPEYHSE-----TPEOMYHNIN-----	1020
Qy	217	ARQETPNEFARBEKFOVSVKAAEDMGTDSDLMFGYTOOSHMOJTEFGKNSRPERVHD	276
Db	103	---YORNEF-----KFOISFVRPVEFHILMTKGTLYLATYQTDWFOQLYNDQOSAPMRYM	154
Qy	277	KQRELELLQRYSDLPMDGK---RMIGMGAVHNHNSG-ESAKLSRSKNRKYLAAGMRYKN	332
Db	155	FMPELITYUPL-NKRPQGGKIGNFSELMIGMHTSNGVGAGOCYOPBNK---EGNDENO	209
Qy	333	LTVMPRI-----WGRIFKEGSGSOP-----DDNDP	357
Db	210	FPQGPVUJIKDYNQGDVWRMGCSRSVSAGORPVEFLVMEKGLKIMAYWMPYVPDQSNPN	269
Qy	358	ILDYUGGVDFEFL-----OLENKNISGCTVYNNRSGKGLOLDYUVPILKGISGYE	410
Db	270	LIDWGYGNNAKIDYRRGRHNEELDYDLFIOYWKRYD--RNHGAFRLQYTYRINPVGTYA	327
Qy	411	QIFQYGGOSLIDYNNHEATSPGCVGLMLN	437
Db	328	QMFNGYGGDLEYEXDYVFSNRLGVGZIRLN	354

RESULT	10			
Q9ZLX5				
ID	Q9ZLX5	PRELIMINARY;	PRT;	355 AA.
AC	Q9ZLX5;			
DT	01-MAY-1999 (TrEMBLrel. 10, Created)			
DT	01-MAY-1999 (TrEMBLrel. 10, last sequence update)			
DT	01-JUN-2001 (TrEMBLrel. 17, last annotation update)			
DE	PURATIVE PHOSPHOLIPASE A1.			
GN	PIDA OR JHP0451.			
OS	Helicobacter pylori J99 (Campylobacter pylori J99).			
OC	Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;			
OC	Helicobacter.			
OX	NCBI_TaxId=85963;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99120557; PubMed=9923682;			
RA	Alm R.A., Ling L.-S.L., Molr D.T., King B.L., Brown E.D., Dolg P.C.,			
RA	Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,			
RA	Tummino F.J., Caruso A., Ulla-Nickelsen M., Mills D.W., Ives C.,			
RA	Gibson R., Merberg D., Mills S.D., Jlang Q., Taylor D.E., Voyls G.F.,			
RA	Trust T.J.;			
RT	"Genomic sequence comparison of two unrelated isolates of the human			
RT	gastric pathogen Helicobacter pylori.",			
RL	Nature 357:176-180(1999).			
DR	EMBL; AE001479; AAD06029.1; .			
DR	InterPro; IPR003187; PLAI.			
DR	Pfam; PF02253; PLAI. 1.			
DR	Complete proteome.			
Q0	SEQUENCE 355 AA; 42329 MW; B3CE9810EBA2FDC9 CRC64;			

Query Match	10.2%;	Score 240.5;	DB 2;	Length 355;
Best Local Similarity	25.1%;	Pred. No. 1e-12;		
Matches	83;	Conservative	44;	Mismatches 109;
				Indels 95;
				Gaps 13;

Oy	157	KKDAKOLEYAKOTPEPLSDDDDRNNTPLMSSRPNNPVLPLPFIHGKPRNSNPNSHE	216
Db	69	KKYLMMDYLCTEYELPEYHNS-----TPEOMYHNHPIN-----	1020
Oy	217	AROFPPNEFARPELKFQVSUVUKAABDAGSDSLMEFGUOOSMOLEFGKNSRPFVND	2767
Db	103	---YORNEF-----KFQISEFVPUVEFRNLTMTKGLYLATYOTNFEQLYNDPOSAFPMIN	154
Oy	277	YOPETELFOPVUSDLPWBGK*-RMIGMGAUHNHSG-ESAKLSRSNNR-----	321
Db	155	FWPELTYUPPL-NKRPFGSKIGNSEIEMIGONHNSNGVGAOCQVRPNKEGNRENDPFGQ	213
Oy	322	-----AYUMAG-----MEW*-NLTVAPRIWGRIFKEGSGSDPD-	354
Db	214	PVUYWDYNGGQKDRVHGGGSRVSAAGALCEFLWMEKGGKILVAAWPUV-----PYDQ	265
Oy	335	-NPDLIDYUGXGVDFEL-----OLENKSINISGTVAIYNDRSGKALDLDYUPLSGI	406
Db	266	SNPOLIDYMGXNAKIDRYRGRRHNFELDLDIFIQWRYD--RWNGAFRLGUTYVRIWPU	3223
Oy	407	SGYRPIFGQSGSLDIYUHNHETSGVGLMLN	437
Db	324	GIYAOFMFGXGDLXUEYDFENSRJGUGIRLN	354

RESULT	11			
ID	Q9XB53	PRELIMINARY:	PRF:	278 AA.
AC	Q9XB53;			
DT	01-NOV-1999 (TrEMBLrel. 12, Created)			
DT	01-NOV-1999 (TrEMBLrel. 12, Last sequence update)			
DT	01-NOV-1999 (TrEMBLrel. 12, Last annotation update)			
DE	KDUI.			
OS	<i>Erwinia carotovora</i> .			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Pectobacterium.			
OX	NCBI_TaxID=554;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 39048, GS101;			
RX	MEDLINE=8065591; PubMed=9402024;			
RA	McGowan S.J., Sebahia M., O'Leary S., Hardie K.R., Williams P.,			
RT	Stewart G.S., Bycroft B.W., Salmood G.P.;			
RT	"Analysis of the carpanem gene cluster of <i>Erwinia carotovora</i> ;			
RT	definition of the antibiotic biosynthetic genes and evidence for a			
RT	novel beta-lactam resistance mechanism.";			
RL	Mol. Microbiol. 26:545-556(1997).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 39048, GS101;			
RX	MEDLINE=98276484; PubMed=9614345;			
RA	McGowan S.J., Bycroft B.W., Salmood G.P.;			
RT	"Bacterial production of carbapenems and clavams: evolution of beta-			
RT	lactam antibiotic pathways.";			
RL	Trends Microbiol. 6:203-208(1998).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 39048, GS101;			
RA	McGowan S.J.;			
RT	Submitted (JUN-1999) to the EMBL/GenBank/DBD databases.			
RT	EMBL: U17224; AAD38237.1; -;			
SO	SEQUENCE 278 AA: 31341 MW; 63769P4F3550E1B7 CRC64;			

Query Match	5.28;	Score 123;	DB 2;	Length 278;
Best Local Similarity	21.68;	Pred. No. 0.012;		
Matches 74;	Conservative 52;	Mismatches 128;	Indels 88;	Gaps 16;

```

QY 51 IDVGATGASTDTANPLDEHEPELTYTALENKTMILNCNALNODIMRLACYDILVHGFT 110
DB 1 MDVRSVSHS-----EHAKTLDITELRKFLIEQIFETPNQYTMYSIDRLIVG-- 48
QY 111 PAVITKRSIRLDETIMO-----TIKGPQVYQETDPTILMNGEKM 154
DB 49 -GIMVDEITFDDIGIGQFVNYFLERRELGLINIGGPAKIVIGTS--YEVNEEEL 104
QY 155 LTKKDAKOLEYAAKQFTPLSLFDDLRNNTPLMSSRPHNPMVYLPIFMHGKPNRSPTPS 214
DB 105 YVGKAKALAFS-----SLDSAKPAKLYNSAPAHAVPTPTIITODDAIKAPLDV 155
QY 215 HEARQFTNEFRAPLKFQVSVKRAAEDLMGTSDLMFGYT--QOSHMOIFNGKNSRP 271
DB 156 KTCNKRITICKYLVPEV-----VETCQ-----LSMGLTRLAEGSNW-----NSMP 194
QY 272 FRVHDYQPEIFLTQPVYSDLPMDGKVRMIGCAVHNS-----NGESAKISRSNRAVYLM 326
DB 195 THTHRRRMEVYF---YEDMAEDTIIIFHM-MGEPHETRLVNHNEQAVISPSMS---IHT 246
QY 327 GMEWKNLTVMPRIMGRIFEGSGSQPDNDPILDYGYGQDVR 368
DB 247 GVGTKNYAF---INGMT---GENLTFDD---MDHIMLMDLR 278

RESULT 12
C000443 PRELIMINARY: PRT: 1686 AA.
AC 000443:
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PHOSPHOINOSITIDE 3-KINASE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=97479209; PubMed=9337861;
RX SEQUENCE FROM N.A.
RA Dorn J., Pages F., Vollma S., Rittenhouse S.E., Zvelebil M.J.,
RA Stein R.C., Waterfield M.D.;
RT "Cloning of a human phosphoinositide 3-kinase with a C2 domain that
RT displays reduced sensitivity to the inhibitor wortmannin.";
RL Biochem. J. 326:139-147(1997).
DR EMBL: Y13367; CAA73797.1; -
DR HSP: P21707; IRYN.
DR InterPro: IPR000008; C2.
DR InterPro: IPR001263; PI3Ka.
DR InterPro: IPR002420; PI3K_C2.
DR InterPro: IPR000341; PI3K_ras_bind.
DR InterPro: IPR000403; PI3_P14_kinase.
DR InterPro: IPR001683; PX.
DR Pfam: PF00168; C2; 1.
DR Pfam: PF00613; PI3Ka; 1.
DR Pfam: PF00792; PI3K_C2; 1.
DR Pfam: PF00794; PI3K_rbd; 1.
DR Pfam: PF00454; PI3_P14_kinase; 1.
DR Pfam: PF00787; PX; 1.
DR SMART: SM00239; C2; 2.
DR SMART: SM00145; PI3Ka; 1.
DR SMART: SM00146; PI3K; 1.
DR SMART: SM00142; PI3K_C2; 1.
DR SMART: SM00144; PI3K_rbd; 1.
DR SMART: SM00312; PX; 1.
DR PROSITE: PS00004; C2_DOMAIN_2; 1.
DR PROSITE: PS00915; PI3_4_KINASE_1; 1.
DR PROSITE: PS00916; PI3_4_KINASE_2; UNKNOWN_1.
DR PROSITE: PS00290; PI3_4_KINASE_3; 1.
SQ SEQUENCE 1686 AA; 190736 MW; E9311C803025C96F CRC64;

```

```

Query Match 4.8%; Score 114; DB 4; Length 1686;
Best Local Similarity 19.0%; Pred. No. 0.92;
Matches 85; Conservative 68; Mismatches 154; Indels 140; Gaps 21;

QY 45 QDNELPIDV-----QSAQGSASTDTANPLDE-----HEPELTATLENKTMILN----- 88
DB 522 EDETPVDLNTKLKLYOIEKPCKEAMTRHPVEELDSYHNOVELALQIENQHRADVITAV 581
QY 89 ---CSALNODIMRLACYDYLIGETPAVITKRSIRLDETIMOITKGRQVYQETDPTI 145
DB 582 RKICSAID-GVETLAI-----TESVKKLRAVNLPRS-----KTADVT 618
QY 146 FLMGNEKGMLTKKDAKOLEYAAKQFTPLSLFDDLRNNTPLMSSRPHNPMVY-----LP 199
DB 619 SLFGGE-----DTSRSSP-RGSLAPENQVYSINQLTAA 651
QY 200 IF-----MHGKPNRSPTPSHEARQFTNEFRAPLKFQVSVKRAAEDL---WGTDSDLM 252
DB 652 IYDLRLHANSGRSPDPCQSSKSVKEMATTQEQLTIF---AAHGSSMWVSNYEXY 707
QY 253 FGYTQOSHMOIFNGKN-SHPF---RVHDYQPEIFLTQ-----PV-YSDLPMDGKVRMI 300
DB 708 YLIGSLSH-----NGKDLFRPIQSKKVGTYKNFFYLKMDLIIPIQISQPLFESVHLHT 763
QY 301 GMGAVHNSGESAKLSRSVN-----RAYMAGMEKNLTVMPRIMGRITPK 345
DB 764 LFGLNQSSGSSPDSSNKKRKGPEALGVKSLPLCDFRRFLTCG-----TKLLYLW----- 812
QY 346 ESGSGQPDNDPILDYGYGQDVRPLYOLENKNISGTVVYNRSGALQDLVYVPLGK 405
DB 813 --TSHTNSVPGTVTKKGVMERYIVQVDFPSPADIIYTTQVDRSIIQOHNLLETEND 870
QY 406 ISGYFOIFQYQSLIDYVNHNEATSEGV 432
DB 871 IKG-----KLLDLHKDSISGL 887

RESULT 13
C063485 PRELIMINARY: PRT: 602 AA.
AC 063485:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE C-RAF ACTIVATED ONCOGENE FUSION PROTEIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87172791; PubMed=3550433;
RA Ishikawa F., Takaku F., Nagao M., Sugimura T.;
RA "Rat c-raf oncogene activation by a rearrangement that produces a
RT fused protein.";
RL Mol. Cell. Biol. 7:1226-1232(1987).
CC -I- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: M15428; AAA42002.1; -
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_kin_actstle.
DR Pfam: PF00069; pkinase; 1.
DR SMART: SM00221; STYK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; kinase; Oncogene; Serine/threonine-protein kinase;
KW transferase.
SQ SEQUENCE 602 AA; 69209 MW; 88B7BFA90FFB02AC CRC64;

```

```

Query Match 4.7%; Score 110.5; DB 11; Length 602;
Best Local Similarity 19.5%; Pred. No. 0.43;
Matches 93; Conservative 68; Mismatches 171; Indels 145; Gaps 23;

```

```

0Y 29 NPVAFVDEVRSENDLQGDNEPRIDVQASATQASPDTPANRPDENHEPRLETYALEKTYLIN 88
Db 87 NOVAVULEKKELETAODRNLGT-----OSQFRAKEEELAEKRRDLRT--NERLSQE 137
0Y 89 CSALNODIMRIACXDYLTHGCTPRAVITKRSIRIDENIWO--TIKGRPOVUOE----- 140
Db 138 VEYILEDOVKRL--NEKLKESMT--TGELIQIKDELQASDVUYKYAREKRLDEQKELIHN 192
0Y 141 -----TDRIFLMGNEK-----MLTKDAKOLEYAAKORTPLSLSPDL 180
Db 193 QNSWMLNTELKTKTELLALGRENELTELKCTLENKEEDALISHEASPSALYS--SSP 251
0Y 181 RNNTPLMSRRHPNRYULYRPMHCKPKNRSPNTPHEKROTPNEFRDELKFOYSUYKA 240
Db 252 NNLSPSTGSOQKTP-----VPADREBARPSGQEKNIIRPKOROSSYWEIEASEVM 304
0Y 241 AEDLMGTDSDLMFGYTOOSH-----QIFNGKNSRPFVHDXOPEI----- 281
Db 305 LSTRIGSGS---FCTVYUGKWHGDNVAKYLKVDPRTPEQLOAFNEVEVALKTRHNVIL 361
0Y 282 -----FLQO-----PVYBDL--PMOSKVMY-----GMAVHNHNSESA 313
Db 362 FMGYTKDKNLAIYVQKCGSSLYNHLNHOEKKFQFMOFLIDIRATQAGMDYLAKKNIHR 421
0Y 314 KLSRSMNRAVYLMAGMEMK-----NLTVPRIRGIRFFKEGSGSQ----- 351
Db 422 DMKS--NNIFLHEGLTVIGDFGLATVKSRL-----SSGQVEOPTGSVLMAPAEVIR 472
0Y 352 -PDDNPDL--DYGQGVDRFLYOLEKNSISGIVTKRNPBSGKALDLDVYPLGKG 405
Db 473 MODNNPESFGSDVSYGIV--LYEL-----MTGLPYSHINNROI-----IFWNG 518

```

```

RESULT      14
059241      PRELIMINARY;          PRT;          821 AA.
AC          059241;
DT          01-NOV-1996 (TREMBLrel. 01, Created)
DT          01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT          01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE          ENO-1,4-BETA-GLUCANASE (EC 3.2.1.4) (CELLULASE) (ENDOGLUCANASE)
DE          (CARBOXYMETHYL CELLULASE) .
OS          Bacillus sp.
OC          Bacteria; Firmicutes; Bacillus/Clostridium group;
OC          Bacillus/Staphylococcus group; Bacillus.
OX          NCBI_TaxID=1409;

RM          [1]
RP          SEQUENCE FROM N.A.
RC          STRAIN=KSM-64;
RA          Sumitomo N., Ozaki K., Ito S.;
RL          Submitted (MAY-1993) to the EMBL/GenBank/DDBI databases.
RC          -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
CC          LINKAGES IN CELLULOSE.
CC          EMBL; M84963; AAA73189.1; .
DR          HSSP; O85465; 1A3H.
DR          InterPro: IPR001547; Glyco_hydro_F5.
DR          Pfam: PF00150; cellulase.1
DR          PROSITE; PS00655; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
KW          Hydrolase, Glycosidase.
SQ          SEQUENCE      821 AA;  90910 MW;  73D438FEF0B4DB5C CRC64;

```

Query Match	4.6%;	Score 109.5;	DB 2;	Length 821;
Best Local Similarity	19.8%;	Pred. No. 0.81;		
Matches 95;	Conservative 54;	Mismatches 163;	Indels 169;	Gaps 25

Oy 9 TLTSLSCALLAIOAKVPVPAFVDEVREND---LGDONELPIDVQNSATGSADTFA 65
| : : | : : | : : | : : |
Db 6 TKOLISLTILIVL-LTLSEPTALAAEGNTRDNFKHLGSDNVKRPSEAGALQLQEYDQG 64
| : : | : : | : : | : : |
Oy 66 NPL-DEHEPELY-----TTALEKNTMLINCAL-----NODIMRLACY-DIIVHGEP 111
| : : | : : | : : | : : |

```

Db      65  MTLVDQHEKIQLGKSMSTHGLQWPEPILINDANAYALANDMESNNIRLAMYVGEAGYASNP 124
Qy      112  AVIKTKRSIRLDEYI-----WQT-----IKGPPQVY 138
           : : : | | |
Db      125  ELIKSRVYKIGIDLAIENDMYIYVWHNHAQRPDPYACAGEDFRRIALAYRNNPILIIY 184
           : : : | | |
Qy      139  QETTD-----IFLMGNEKMLTKDKAQLEYAAKQTPPLSLSPDLBN---NTPLW 187
           : : : | | : | | | | | |
Db      185  ELANFPSSNNNGAGIIPNNEGWMAYK-----EXADPIVLELRSGNADNDIIVGSPMW 239
           : : : | | : | | | | | |
Qy      188  SSRP-----HNPYVLPIF--MHGKPNRS--PNTDSHEARQTPNEFRAPELKFQ 233
           : | | : | | : | | | | | |
Db      240  SQRPDLADNPIIDHHFTWYVHFTGSHAASTESEYPRETPNSEKGVMSNTRYA----LE 295
           : | | : | | : | | | | | |
Qy      234  VSVKYYKAAEDIMGT-----DSDIMEGTQQ-----SHWQIFNGKNS-----RP 271
           : | | : | | : | | : | | |
Db      296  NGVAVFATE--MGTSGAAGDGPFEDEADWVIEFLINENNISMAMWSLITN--KNEVSGATP 352
           : | | : | | : | | : | | |
Qy      272  FRVH-----DTPLEFLTPQYISDLPM--D 294
           : | | : | | : | | : | | |
Db      353  FELGKSNATSLDPGPDQVWPVEELSLGSEYVRARIKGVNEP--IDRTQYTYKLMDFND 409
           : | | : | | : | | : | | |
Qy      295  GKVRKIGGAVHHSHGES-----AKTSRSMNNRAYILMAGEMKNLTVMPIRWG 341
           : | | : | | : | | : | | |
Db      410  GTKOGFV-----NGDSPVEDVYIENEAGALKLSGLDASNDVSEGVYMANARLSADGWG 463
           : | | : | | : | | : | | |
Qy      342  R 342
           : | | : | | : | | : | | |
Db      464  K 464
           : | | : | | : | | : | | |

```

```

RESULT      15
Q9F216      Q9F216      PRELIMINARY:      PRT:      824 AA.
AC          Q9F216;
DT          01-MAR-2001 (TrEMBLrel. 16, Created)
DT          01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT          01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE          CELLULOSE.
OS          Bacillus sp.
OC          Bacteria; Firmicutes; Bacillus/Clostridium group;
OC          Bacillus/staphylococcus group; Bacillus.
OX          NCBI_TaxID=1409;
RN          [1]
RP          SEQUENCE FROM N.A.
RC          STRAIN=KSM-S237;
RX          MEDLINE=21036886; PubMed=11193393;
RA          Hakemada Y., Hatada Y., Koike K., Yoshimatsu T., Kawai S.,
RA          Kobayashi T., Ito S.;
RT          "Deduced amino acid sequence and possible catalytic residues of a
RT          thermostable, alkaline cellulase from an alkaliphilic Bacillus
RT          strain.";
RL          Biosci. Biotechnol. Biochem. 64:2281-2289(2000).
DR          EMBL: AB018420; BAB19360.1; -.
DR          InterPro: IPR001547; Glyco_hydroc.F5.
DR          Pfam: PF00150; cellulase.1.
DR          PROSITE: PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
SQ          SEQUENCE      824 AA: 91564 MW: 65FA940FELID729B9 CRC64;

```

Query Match 4.6%; Score 108; DB 2; Length 824;
Best Local Similarity 19.8%; Pred. No. 1.1;
Matches 95; Conservative 60; Mismatches 159; Indels 166; Gaps 28;

```

OY 9 TILSISPAIIAQAKAVPVPAVFVEBSEEND--LGGDNELPIDVOSATOSASTDA 65
    | : : | | : | | : | | : | | : | | : | | : | | : | | : | | : |
Db 7 TKQILSILLILVL-LLSIFPALAAEGNTRREDNKHLGLSNDNVKRPSEAGALQLEVDGQ 65
    | : : | | : | | : | | : | | : | | : | | : | | : | | : | | : |
OY 66 NPL-DEHEPELY-----TTALENKTMLINCAL-----NODIMRLACTDYLTHGE-- 109
    | : : | | : | | : | | : | | : | | : | | : | | : | | : | | : |
Db 66 MTLVQDQHEKIQIKRGMSYHGQMFPEELLINDAKYALSDWDSNMIRLAMY--VGGNGY 121
    | : : | | : | | : | | : | | : | | : | | : | | : | | : | | : |
OY 110 --TPAVIKTK-----RSIRLDEFT--WQF-----IKGP 134
    | : : | | : | | : | | : | | : | | : | | : | | : | | : | | : |

```

```
Db 122 ATNPBLIKORVIDGIELALENDMYVIVDWHVHAPGDPDPVYAGAKDFPREIALALYPNNP 181
Qy 135 QVVOETTP-----IFLGNKMGMLTKDAKOLEYAKOFTPLSLFEDLDRN-----N 183
Db 182 HITTELANEPSNNNGAGIPNNEGMAK-----ETADPIYEMLRKSGNADDNIIIVG 236
Qy 184 TPLWSSRP-----HNPVYLPFIMHGKPNRSPNTPSHEARQFTPNFRAPELK-- 231.
Db 237 SPNMSORPDLAADNPIDHHTMYTVHFYTGSHAASSTESYPS-----TPNSERGNVMSNT 291
Qy 232 ---FOVSVKYKAAEDLMGT-----DSDLMFCTTQQ-----SHWQIFNGKNS-- 269
Db 292 RYALENGAVAFATE--WGTSQASGDGPFYDEADWIEFLNENNISMAMNSILTN--KNEYS 348
Qy 270 ---RPPRV-----HDYOP--EFLT-----OPV-----YSDLPW-- 293
Db 349 GAFTEPELCKSNATNIDPDPDHVMAPEELSLSGEYVRARIKGVNTEPIDRTKTYKWLMDP 408
Qy 294 -DGKVRMIGMGAVHHSNGESAKLSRSWNRAYLWAGME-----WKNLTVMPRIWGR 342
Db 409 NDGTQO--GEGVNSDSPNKELIAVDNENNTLKVSGLDVSNDSGDNFMANARLSANGWGK 466
```

Search completed: November 30, 2001, 14:26:28
Job time: 565 sec


```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
CC -----
CC DR EMBL: X76902; CAAS4224.1; -.
CC DR PIR: C36971; C36971.
CC DR PIR: S40130; S40130.
CC DR InterPro: IPR003187; PLAL1.
CC DR Pfam: PF02253; PLAL1.
CC KM HydroLase; Lipid degradation; Outer membrane; Signal; Calcium.
CC FT SIGNAL 1 20 BY SIMILARITY.
CC FT CHAIN 21 289 PHOSPHOLIPASE A1.
CC FT ACT_SITE 164 164 BY SIMILARITY.
CC SQ SEQUENCE 289 AA; 32944 MW; D7516CFBBD406997 CRC64;
CC
CC Query Match 13.6%; Score 321.5; DB 1; Length 289;
CC Best Local Similarity 33.3%; Pred. No. 8.3e-19;
CC Matches 87; Conservative 42; Mismatches 115; Indels 17; Gaps 7;
CC
CC QY 179 LDRNRTPLMSSRPHNPMLVPLIFPMHGKRNSEPTNFSHARQPTNPFNAPELKFOVSXV 238
CC I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I ::
CC 43 LQEHNP-FTLYPYSENVLALTY-----TSDLNKAISYNNSDNA-NKDEKFKPQLSLAF 95
CC
CC QY 239 KAAEDLWGTDSDLMFGYQQOQHMOIFNGKNSRPFVNDYQEIFLTPY-VSDLFMDGV 297
CC I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I ::
CC 96 PLMRCILTDNSILGASYSYQSRWMOQLSNTGSAFPFRETVEPOLFLGFATDYSVGDW--TL 153
CC
CC QY 298 RMIGCAVHSHSGESAKLSRSNRYRYLMAGMWKRKLTYPRWGRIFKRGSSQDDND 357
CC I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I ::
CC 154 RDAEFGYHSHQNSGRSDPTSRSMNRLYSRLMAONGMLVEKFWYI-----GDTSDKN 207
CC
CC QY 358 ILDYGYGDVRELYOLEKSNISGTVRYPNSRGALDIDYVYPLGKISGYFOIGYG 417
CC I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I ::
CC Db 208 ITRYGYQYQLKIGYLD-GEAVLSAKGQYNNMNTGYGAEIGVSYPLTKHVRFTYQYSGIG 266
CC
CC QY 418 OSLLIDYNEHATSFYGYGLMND 438
CC I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I ::
CC Db 267 ESLIDYDFNQTRFVGMVLMND 287
CC
CC RESULT 5
CC GUN_BACSI STANDARD: PRT: 800 AA.
CC ID_GUN_BACSI P06564;
CC AC P06564;
CC DT 01-JAN-1988 (Rel. 06, Created)
CC DT 01-JAN-1988 (Rel. 06, Last sequence update)
CC DT 15-JUL-1999 (Rel. 38, Last annotation update)
CC DE ENDOGUCANASE PRECURSOR (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE)
CC OS (ALKALINE CELLULOSE).
CC OS Bacillus sp. (strain 1139).
CC OS Bacteria; Firmicutes; Bacillus/clostridium group;
CC OC Bacillus/staphylococcus group; Bacillus.
CC NCBI_TaxID=1411;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=87085443; PubMed=3098909;
CC RA Fukumori F., Kudo T., Narahashi Y., Horikoshi K.;
CC RT "Molecular cloning and nucleotide sequence of the alkaline cellulase
CC gene from the alkalophilic Bacillus sp. strain 1139."
CC RL J. Gen. Microbiol. 132:2329-2335(1986).
CC -1 CATALYTIC ACTIVITY: ENDOMYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
CC LINKAGES IN CELLULOSE.
CC -1 MISCELLANEOUS: ALKALOPHILIC BACILLUS SP., STRAIN 1139, PRODUCES
CC ONE ALKALINE CELLULOSE (PH OPTIMUM 9.0). THIS STRAIN IS NOT A TRUE
CC CELLULOTIC MICRO-ORGANISM BECAUSE THE ENZYME IS UNABLE TO
CC HYDROLYSE NATIVE CELLULOSE.
CC -1 SIMILARITY: BELONGS TO CELLULOSE FAMILY A (FAMILY 5 OF GLYCOSYL
CC HYDROLASES).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```

CC
CC EMBL: M15743; AAA22305.1; -.
DR DR EMBL: D000066; BAA00045.1; -.
DR DR PIR: A29003; A29003.
DR InterPro: IPR001547; Glyco_hydro_F5.
DR Pfam: PF00150; cellulase; 1.
DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5_1.
KM Cellulose degradation; Hydrolase; Glycosidase; Signal.
FT SIGNAL 1 30
FT CHAIN 31 800
FT ACT_SITE 190 190
FT ACT_SITE 305 305
SQ SEQUENCE 800 AA; 88602 MW; 70CADD7B6DAD55CF CRC64;

Query Match 4.6%; Score 109.5; DB 1; Length 800;
Best Local Similarity 19.8%; Pred. No. 0.47;
Matches 95; Conservative 54; Mismatches 163; Indels 169; Gaps 25;

QY 9 TLTSLSCALLAIGQAKAVPNPVAFVDEVRSND---LGDNELPIVQASQASSTGTA 65
   |:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 7 TKDILSILILVL-LLSLPTPLAAEGNTRBDNFKHLGNDNVKRPSEAGALQDEVGQ 65

QY 66 NPL-DEHEPELY-----PTALENKTMLNCSAL-----NODMIRLACY-DLIVHGEP 111
   |:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 66 MTLVDQHEKQLQLGKMSHGLQMPPELLNDNAYKALANDWESNMIRLAMYGENGYASNP 125

QY 112 AVIKTKRSIRLDET-----WQT-----IKGRQVYV 138
   |:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 126 ELKSRVLIKIDLAIEENDMYIVDMHVAHAGDPDPVYAGAEDFERDIALYRNPPIIY 185

QY 139 QETDP-----IFLMGNEKGMILTKKAKOLEYAAKQFTPLISFDDRN---NTPLW 187
   |:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 186 ELANPEPSSNNNGAGIPNNEGWMNAVK-----EYADPIVEMLRSGNADNDIIIVSPNW 240

QY 188 SSRP-----HNPPYVLPIF--MGKPNRS--PPTPSHEARQFPNPEFAPELKQ 233
   |:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 241 SQRPLDADNPIDDHHTMYTFHYTGSHAASTESTPPEPENSEKGNVMSNTRYA---LE 296

QY 234 VSVYVKAEDLMGT-----DSDLFGGYTQ-----SHMOIFNGKNS-----RP 271
   |:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 297 NGVAVFATE--WGTSGAANGDGGPYRDEADVWIEFLENNENNISMANNSLNL-KNEVSGATFP 353

QY 272 FRVH-----DYOPELFLTPVYSDLFW---D 294
   |:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 354 FELGKSNATSLDPGDQVWPVEELSLGEVYRARIKGVNVEP---IDRTKTKYKVLMDPND 410

QY 295 GKVRKIGGAVHHSGES-----AKISRSNNRATYLAGMGMKMLTVPRFNG 341
   |:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 411 GTKGQGFV-----NGDSPVEDVYIENAGALKLSGLDASNDVSEGNWYANARLSADGNG 464

QY 342 R 342
   |:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 465 K 465

RESULT 6
AMD_YEAST ID AMD_YEAST STANDARD; PRT; 810 AA.
AC P15274;

DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE AMP DEAMINASE (EC 3.5.4.6) (WYOADENYLATE DEAMINASE).
GN AM1 OR AMD OR YML035C.
OS Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
```

OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90105403; PubMed=2690949;
 RA Meyer S.L., Kvalnes-Krick K.L., Schramm V.L.;
 RT "Characterization of AMP, the AMP deaminase gene in yeast. Production
 of amd strain, cloning, nucleotide sequence, and properties of the
 RT protein."
 RL Biochemistry 28:8734-8743(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288C / AB972;
 RA Badcock K., Churcher C., Barrell B.G., Rajandream M.A., Walsh S.V.;
 RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY
 CC METABOLISM.
 CC -1- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).
 CC -1- PATHWAY: PURINE NUCLEOTIDE CYCLE.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SIMILARITY: BELONGS TO THE ADENOSINE AND AMP DEAMINASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: M30449; AAA34420.1; -
 CC EMBL: Z46659; CAA86620.1; -
 CC PIR: A33365; A33365.
 CC SGI: S0004498; AMD1.
 CC InterPro: IPR001365; A_deaminase.
 CC Pfam: PF00962; A_deaminase.1.
 CC PROSITE: PS00485; A_DEAMINASE.1.
 CC HydroLase: Nucleotide metabolism.
 CC FT ACT_SITE 422 422 POTENTIAL.
 CC FT ACT_SITE 631 631 POTENTIAL.
 CC FT ACT_SITE 707 707 POTENTIAL.
 CC FT ACT_SITE 708 708 POTENTIAL.
 CC FT CONFLICT 568 568 F -> C (IN REF. 1).
 CC SC SEQUENCE 810 AA; 93301 MW; 7A6DCB439B45C93 CRC64;

 Query Match 4.2%; Score 98.5; DB 1; Length 810;
 Best Local Similarity 21.3%; Pred. No. 3.7;
 Matches 74; Conservative 39; Mismatches 132; Indels 103; Gaps 15;

RESULT 7
 ID BCK1_YEAST STANDARD; PRT: 1478 AA.
 AC 001389; P32894;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE SERINE/THREONINE PROTEIN KINASE BCK1/SLK1/SSP31 (EC 2.7.-.-).
 GN BCK1 OR SLK1 OR SSP31 OR LAS3 OR YOL095W OR J0906.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92186847; PubMed=1545797;
 RA Costigan C., Gehring S., Snyder M.;
 RT "A synthetic lethal screen identifies SLK1, a novel protein kinase
 RT homolog implicated in yeast cell morphogenesis and cell growth."
 RL Mol. Cell. Biol. 12:1162-1178(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92104496; PubMed=1840547;
 RA Irie K., Araki H., Oshima Y.;
 RT "A new protein kinase, SSP31, modulating the SMP3 gene-product
 RT involved in plasmid maintenance in Saccharomyces cerevisiae."
 RL Gene 108:139-144(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EG123;
 RX MEDLINE=92107166; PubMed=1729597;
 RA Lee K.S., Levin D.E.;
 RT "Dominant mutations in a gene encoding a putative protein kinase
 RT (BCK1) bypass the requirement for a Saccharomyces cerevisiae protein
 RT kinase C homolog."
 RL Mol. Cell. Biol. 12:172-182(1992).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288C;
 RX MEDLINE=95116706; PubMed=7871887;
 RA Miosga T., Boles E., Schaeff-Gerstenschlaeger I., Schmitt S.,
 RA Zimmermann F.K.;
 RT "Sequence and function analysis of a 9.74 kb fragment of
 RT Saccharomyces cerevisiae chromosome X including the BCK1 gene."
 RL Yeast 10:1481-1488(1994).
 RN [5]
 RP SEQUENCE OF 602-1104 FROM N.A.
 RA Cusick M.E.;
 RL Submitted (XX-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: SERINE/THREONINE PROTEIN KINASE INVOLVED IN A SIGNAL
 CC TRANSDUCTION PATHWAY THAT PLAY A ROLE IN YEAST CELL MORPHOGENESIS
 CC AND CELL GROWTH. THIS PATHWAY SEEMS TO STARTS BY SMP3; THEN
 CC INVOLVE THE KINASE PC1 THAT MAY ACT ON THIS KINASE. BCK1 PROBABLY
 CC PHOSPHORYLATES MK1 AND MK2 WHICH THEMSELVES PHOSPHORYLATE THE
 CC MPX KINASE.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE KINASE KINASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: M84389; -; NOT_ANNOTATED_CDS.
 CC EMBL: D10389; BAA01226.1; -
 CC EMBL: X60227; CAA42788.1; -
 CC EMBL: X77923; CAA54896.1; -
 CC EMBL: Z49370; CAA89389.1; -

```

DR EMBL; Z49369; CA89388.1; -.
DR EMBL; M88604; AAA21179.1; -.
DR PIR; S20117; S20117.
DR PIR; S22285; S22285.
DR PIR; J01118; J01118.
DR PIR; J01432; J01432.
DR HSP; P24941; IAO1.
DR SGD; S0003631; BCK1.
DR InterPro; IPR000719; Euk_Pkinase.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
KW ATP-binding; Phosphorylation.
FT DOMAIN 1175 1440 PROTEIN KINASE.
FT NP_BIND 1181 1189 ATP (BY SIMILARITY).
FT BINDING 1204 1204 ATP (BY SIMILARITY).
FT ACT_SITE 1303 1303 BY SIMILARITY.
FT MOD_RES 1134 1134 PHOSPHORYLATION (BY PKC) (POTENTIAL).
FT VARIANT 1119 1119 T -> P (IN BCK1-19; ACTIVATION).
FT VARIANT 1120 1120 I -> K (IN BCK1-11; ACTIVATION).
FT VARIANT 1120 1120 I -> T (IN BCK1-16; ACTIVATION).
FT VARIANT 1146 1146 G -> V (IN BCK1-10; ACTIVATION).
FT VARIANT 1174 1174 A -> P (IN BCK1-20; ACTIVATION).
FT CONFLICT 59 59 E -> I (IN REF. 2).
FT CONFLICT 79 79 E -> V (IN REF. 3).
FT CONFLICT 264 264 A -> P (IN REF. 3).
FT CONFLICT 279 279 N -> I (IN REF. 3).
FT CONFLICT 703 714 RYPTSTYDTR -> STPKRVTMT (IN REF. 3).
FT CONFLICT 795 795 S -> A (IN REF. 3).
FT CONFLICT 802 802 L -> V (IN REF. 3).
FT CONFLICT 808 808 A -> S (IN REF. 3).
FT CONFLICT 903 903 T -> N (IN REF. 3).
FT CONFLICT 919 919 T -> N (IN REF. 3).
FT CONFLICT 960 962 ADA -> RDR (IN REF. 5).
FT CONFLICT 1086 1104 RPPVDDSSYEPIQGLNGK -> VPIAHTSSYRMDLTGVKN
H (IN REF. 5).
SQ SEQUENCE 1478 AA; 164194 MW; D586C3A497A5BB3 CRC64;
Query Match 4.1%; Score 97.5; DB 1; Length 1478;
Best Local Similarity 23.0%; Pred. No. 10;
Matches 110; Conservative 50; Mismatches 175; Indels 143; Gaps 26;
OY 2 KVSITSLTSL-----ILSCFAILAIQOAKAVPNPVAFVDEVRESENDLGDNDLPT---DV 53
DB 912 KVNNSNSTVSTNSITFTSPSPFLKRGNSKRVVSTSAAD-ITFEENDITPADPPEFSDSD 970
OY 54 QSATQASTDTANPLDEHEPELYTTALENK--TMDLINCASALNODIMRLACYDTLVHGETP 111
DB 971 SDDSSSSDDIDMSKKKAPE--TNNENKKDEKSDNSTHSELF-----YDS---QTO 1019
OY 112 AVITKTRIRLDEITWQITKGPQVYIETDPTFLMGNEKGMILTAKAKOLEVAAKQFT 171
DB 1020 DKMERK-----MTFRSPPEVYVO-NLEKFPFRANLDPKPT-----EGIASPTS 1061
OY 172 PLTSLFDDLRNNT-----PLWSSRPNPMVYLPFIMHG-----KPNSSPNP----- 213
DB 1062 PKSLDLSLSPKNVASSRIEPTSPSPRPVPPDSSYEFTIQGLNGKNKPLMOAKTPKTKTIR 1121
OY 214 --SHEARQFTNEFRAPELKQVSVKVAE-DLWGTDSDLWFGYTOOSHMOIFNGKNSR 270
DB 1122 TIAHNASLARKN-----SVKLKRONTKMGT---RMVETENHMSINNAKNSK 1167
OY 271 PPRVHDYPELFTLPQVYSDLPWGKVMIGMGA-----VHNSGESAKISR----- 317
DB 1168 -----GEYKEFAM-WKGMETKGSFGAVYLCLNVTTGEMAMVAKQVEVPKY 1211
OY 318 -SWNRAVILMAGMEKMLVMPRIWGRIFKEGSGSPDNDPDLDDYGGDVAFLQLENK 376

```

```

DB 1212 SSQNEALIS-----TV-----EALRSEVSTLKDDHLNIYVYLCF-----ENK 1249
OY 377 SNISGTAVYNPRSGCALQLDVYVPLGKISGYFOIFOGYQCSLDVYHHEATSEFVGL 434
DB 1250 NNIT-----SIFLEY--AGGSVGSILRMGRFDEPLI--KHLITVYLKGL 1291
RESULT 8
PBPB_STRPN (
ID PBPB_STRPN STANDARD; PRT; 719 AA.
AC 004707;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (PBP-1A) (EXPORTED PROTEIN 2).
GN PONA OR EXP2.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacilllus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
[1]
SEQUENCE FROM N.A.
RX STRAIN=45607, AND 63915;
RC MEDLINE=93010977; PubMed=1396576;
RA Martin C., Sibold C., Hakenbeck R.;
RT "Relatedness of penicillin-binding protein 1a genes from different
RT clones of penicillin-resistant Streptococcus pneumoniae isolated in
RT South Africa and Spain."
RL EMBO J. 11:3831-3836(1992).
[2]
SEQUENCE FROM N.A.
RX MEDLINE=92325042; PubMed=1624444;
RA Martin C., Briese T., Hakenbeck R.;
RT "Nucleotide sequences of genes encoding penicillin-binding proteins
RT from Streptococcus pneumoniae and Streptococcus oralis with high
RT homology to Escherichia coli penicillin-binding proteins 1a and 1b."
RL J. Bacteriol. 174:4517-4523(1992).
[3]
SEQUENCE OF 293-369 FROM N.A.
RX STRAIN=RM6;
RC MEDLINE=95020625; PubMed=7934910;
RA Pearce B.J., Yin Y.B., Masure H.R.;
RT "Genetic identification of exported proteins in Streptococcus
RT pneumoniae."
RL Mol. Microbiol. 9:1037-1050(1993).
CC -1- FUNCTION: CELL WALL FORMATION.
CC -1- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- MISCELLANEOUS: THESE STRAINS ARE PENICILLIN-SENSITIVE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X67873; CAA48073.1; -.
DR EMBL; X67872; CAA48072.1; -.
DR EMBL; M80527; AAA26956.1; -.
DR PIR; S28038; S28038.
DR InterPro; IPR001264; Transglycosyl.
DR InterPro; IPR001460; Transpeptidase.
DR Pfam; PF00912; Transglycosyl; 1.
DR Pfam; PF00905; Transpeptidase; 1.
DR ProDom; PD001895; Transglycosyl; 1.
KW Peptidoglycan synthesis; Antibiotic resistance; Cell wall;
KW Multifunctional enzyme.
FT ACT_SITE 370 370 ACYLATED BY PENICILLIN (BY SIMILARITY).
FT DOMAIN 657 683 SER-RICH.
FT VARIANT 124 124 T -> A (IN STRAIN R6).
FT VARIANT 386 386 V -> I (IN STRAIN 63915).

```

FT VARIANT 388 388 D -> E (IN STRAIN R6).
 FT VARIANT 397 397 E -> K (IN STRAIN 63915).
 FT VARIANT 523 523 M -> I (IN STRAIN 63915).
 FT VARIANT 533 533 E -> D (IN STRAINS 63915 AND R6).
 FT VARIANT 540 540 T -> S (IN STRAINS 63915 AND R6).
 FT VARIANT 657 657 S -> N (IN STRAIN 63915).
 SO SEQUENCE 719 AA; 79745 MW; 5BC6C6A93BFA970A CRC64;

Query Match 4.0%; Score 95; DB 1; Length 719;
 Best Local Similarity 20.1%; Pred. No. 6;
 Matches 96; Conservative 70; Mismatches 196; Indels 116; Gaps 25;

QY 3 VSLSTLTLSI-----LSCAIIAIIQAQKAVPNPVAVDVRENDL 43
 DB 13 LLSISFLVIAIIVLGGVFFYYVSKAPSLSESKEILTYDKNQLIADLGSERRV 72
 QY 44 -GODNELPIDVOSATOS-----ASTDTANPLDEHPELYTALENKTMILNCSALN 93
 DB 73 NQAANDIFPDLYKAIYSIEDHREFDRGIDTITILGAFRLNLSNLSLGG-----STLT 126
 QY 94 QDMLRACYDTLVHGETPVAIKTKRSIRLDETIMQTIKGRQVYVQET---TDPIFLMG 149
 DB 127 QOLIKLFEYESTSDQITIS-----RKAQEAMLAIQLEQKATKQEIILTYINKYMSN 178
 QY 150 NKGMLTK-----KRAQLE-----YAKQFTPLS-LISFDLDRNNTPLMSSRP 191
 DB 179 GNYGMOTAQNTYKGLNLSLPQALLAGMPQAPQDYDPSHPREAQDRRLNVL--SEM 236
 QY 192 HNPWVLYPIFMHGKPRSPPTPSHEAROFPTNPFRAPE-----LKRQVS-VKRYKAEDLM 245
 DB 237 KNGGYI-----SAQYERKAVNTPTDGLQSLKSNTPAYMDNLKEVINVEETGYNL 292
 QY 246 GTDSDLMFEGYTO--QSH-MQIFNGKNSRPRVHYDQPEITLQPVYSDLPMGDKVMIGM 302
 DB 293 TGTMDVYTVNDDQAKHLMWDIYNTDEVAYVPDELDQ-----VASTIVDVS-NGKV-IAQL 345
 QY 303 GAVHNSN-----GESAKLSRSNNRAYLMAGMKMLT-VMPII-----WRIKREKS 348
 DB 346 GAHQSSNVSFGINQAVETNRDM-----GSTMKPITTYAPALEXYDSTATIYHDEP 398
 QY 349 GSQPDNDPILDY-YGY-GDVRFYLOENKSNISGTVRYN-----PRSGKALOLDY 398
 DB 399 YNPGINTYVYNNDRGYFGNITLQYALQOSRNVPAVETLNKVLNRAKFTLNLGLGIDY 456

RESULT 9
 GUN_PAERO STANDARD; PRT: 397 AA.

AC P23548;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DE 15-JUL-1998 (Rel. 36, Last annotation update)
 DE ENDOGLUCANASE (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE) (CELLULASE).
 OS Paenibacillus polymyxa (Bacillus polymyxa).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Paenibacillus.
 OX NCBI_TaxID=1406;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90170877; PubMed=2307659;
 RA Baird S.D., Johnson D.A., Seligy V.L.;
 RT "Molecular cloning, expression, and characterization of
 RT endo-beta-1,4-glucanase genes from Bacillus polymyxa and Bacillus
 RT circularis";
 RL J. Bacteriol. 172:1576-1586(1990).
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
 CC LINKAGES IN CELLULOSE.
 CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY A (FAMILY 5 OF GLYCOSYL
 CC HYDROLASES).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: M33791; AAA22631.1; -.
 DR PIR: A35136; A35136.
 DR HSSP: P54583; 1ECE.
 DR InterPro: IPR001547; Glyco_hydro_F5.
 DR Pfam: PF00150; cellulase_1.
 DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; 1.
 KW Cellulose degradation; Hydrolase; Glycosidase.
 FT ACT_SITE 194 194 PROTON DONOR (BY SIMILARITY).
 FT ACT_SITE 317 317 NUCLEOPHILE (BY SIMILARITY).
 SO SEQUENCE 397 AA; 44357 MW; B9C2E802C04F0A2A CRC64;

Query Match 4.0%; Score 94.5; DB 1; Length 397;
 Best Local Similarity 25.7%; Pred. No. 2.9;
 Matches 46; Conservative 21; Mismatches 61; Indels 51; Gaps 10;

QY 301 GMGAVHNSGESAKLSRSNNRAYLMAGMKMLTVMPIRIGR-----IFKEG----- 347
 DB 43 GNTIVDESGKEAFNGLNW-----FGLERPNPT-LHGLMSRMDMLDYKKEGYLIR 95
 QY 348 -----SGSQPDD-----NPDIIDYGYGDVRFYLOENKSNISGTV-----RYNPRS 389
 DB 96 LPSNQLFDSSSRPDSIDYHKNPDLV---GLNPQLQIDMKLIERACGORGIDIIIDRHPRGS 152
 QY 390 GKALOLDYV--YPLKGIGISGYFOIFGY-----GQSLIDVNHETSQVGLMLNDW 439
 DB 153 G-GQSELWYTSYQPESSWISDMKRLADRYKNPNPTVIGADLHNPFGQASWGTSNASTDW 210

RESULT 10

BAC2_MOUSE STANDARD; PRT: 716 AA.

AC P97303;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE TRANSCRIPTION REGULATOR PROTEIN BACH2 (BFB AND CNC HOMOLOG 2).
 GN BACH2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C;
 RX MEDLINE=97042438; PubMed=8887638;
 RA Oyake T., Itoh K., Motohashi H., Hayashi N., Hoshino H., Nishizawa M.,
 RA Yamamoto M., Igarashi K.;
 RT "Bach proteins belong to a novel family of BFB-basic leucine zipper
 RT transcription factors that interact with Mafk and regulate
 RT transcription through the NF-E2 site";
 RL Mol. Cell. Biol. 16:6083-6095(1996).
 CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS AS REPRESSOR OR
 CC ACTIVATOR. BINDS, IN-VITRO, TO NF-E2 BINDING SITES. PLAY IMPORTANT
 CC ROLES IN COORDINATING TRANSCRIPTION ACTIVATION AND REPRESSION BY
 CC MAFK.
 CC -1- SUBUNIT: HETERODIMER OF BACH2 AND MAFK.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSION RESTRICTED TO MONOCYTES AND
 CC NEURONAL CELLS.
 CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY. CNC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 BFB/POZ DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: D86604; BAI13138.1; -
DR HSSP: P05412; JFOS.
DR MGI: MGI:894679; Bach2.
DR InterPro: IPR000210; BTB_POZ.
DR InterPro: IPR001871; bZIP.
DR Pfam: PF00651; BTB; 1.
DR Pfam: PF00170; bZIP; 1.
DR SMART: SM00338; BRLZ; 1.
DR SMART: SM00225; BTB; 1.
DR PROSITE: PS50097; BTB; 1.
DR PROSITE: PS00036; BZIP_BASIC; 1.
KW Transcription regulation; Activator; Repressor; DNA-binding;
KW Nuclear protein.
FT DOMAIN 37 103 BTB.
FT DNABIND 162 168 POLY-GLU.
FT DNABIND 527 542 BASIC MOTIF.
FT DOMAIN 550 572 LEUCINE-ZIPPER.
SQ SEQUENCE 716 AA; 78935 MW; 913283731AE24333 CRC64;

Query Match 4.0%; Score 94; DB 1; Length 716;

Best Local Similarity 23.1%; Pred. No. 7.2; Mismatches 140; Indels 156; Gaps 24;

Matches 101; Conservative 40; Mismatches 140; Indels 156; Gaps 24;

14 SCFAITAIQAKAVNPVAFV---DEV--RSENDLG-----QDNELPIDVQASQTASAST 62
124 SCFSLTQIOLNREGD--LFVCRKDSACORQEDHNSAGEEHEEETMDSTAMACMT 181
63 D-----TANPLDEH-----EPELYTALENKTMILNCALNQ---DIMRLAC 101
182 DOMLDPISFEATAIPIVAKEEALPESVPTDKENSEK---GALNQYPRKKYQLAC 237
102 ---YDILVHG-----ETPAVTKTKRSIL---DET 125
238 TKNVTSADSHSGFASFPFSESDSPGNSLKPGLPMGOIKSEPPSEETEEESTLCLSGDET 297
126 IWOTIKGRPOVYVQETDPI-----FLMGNEKGMLTRK 158
298 ---DIKDRPGOVEMRKQSPARPTSTRGACADLRSRVSSPSCLRSIFGITKGV---- 350
159 DAKOLEYAKOFTPLSLF-----DIDRNNTPLMWS--RPHNPMVYLPIFMHCK 205
351 ESTGTPSTSOQ--PLVRSACPFNKGISQGDLTDTYPLAGNYGQPHVGOQKDVSNFAMGS 408
206 PNRSP-----NTPSHARQFTPNFRAPELK-----FGYSYKVAEAEDLMTGDSL 251
409 PLRGPGPETLCEFSSSPCSGARFLATEHQEPGLMGDMYNOVROIK--CEOSYGTNSSD 467
252 WFGYTOQSHWOJFNCKNSRPFVNDYOPEIFLQPV--YSDLPMGDKVMIAGAVHHNS 309
468 ESG-----SFSADDESCHVQORGOEVKLPFPVDQITDLPRNDFOQMIMK---HKL 516
310 GESAKL-----SRSMNR 321
517 SEOLEFIHDIRRRSKNR 533

RESULT 11
DP3A_BACSU STANDARD: PRG: 1115 AA.
AC 034623;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE DNA POLYMERASE III ALPHA SUBUNIT (EC 2.7.7.7).
GN DNAME.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/clostridium group;
OC Bacillus/staphylococcus group; Bacillus.

OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98048467; PubMed=9387221;
RX Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
RT "Sequencing and functional annotation of the Bacillus subtilis genes
in the 200 kb trnB-dnaB region";
RL Microbiology 143:3431-3441(1997).
CC -1- FUNCTION: DNA POLYMERASE III IS A COMPLEX, MULTICHAIN ENZYME
RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN BACTERIA.
CC THIS DNA-POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
CC THE ALPHA CHAIN IS THE DNA POLYMERASE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: N-DEOXYNUCLEOSIDE TRIPHOSPHATE -
CC N PYROPHOSPHATE + DNA(N).
CC -1- SUBUNIT: CONTAINS A CORE (COMPOSED OF ALPHA, EPSILON, AND THETA
CC CHAINS) THAT ASSOCIATES WITH A TAU SUBUNIT WHICH ALLOW THE CORE
CC DIMERIZATION TO FORM THE POLIIT' COMPLEX. POLIIT' ASSOCIATES WITH
CC THE GAMMA COMPLEX (COMPOSED OF CHAINS GAMMA, DELTA, DELTA', PSI,
CC AND CHI) AND WITH THE BETA CHAIN (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-C FAMILY. DNAME
CC SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL: AF008220; AAC00338.1; -
DR EMBL: Z99118; CAB14883.1; -
DR Subtilist: BG12583; dnaE.
DR InterPro: IPR003141; PHP_N.
DR InterPro: IPR002309; tRNA-synt-2.
DR Pfam: PF02231; PHP_N; 1.
DR Pfam: PF01336; tRNA_ant1; 1.
DR Pfam: PF00481; POLIITAC; 1.
DR SMART: SM00481; POLIITAC; 1.
KW Transferrase; DNA-directed DNA polymerase; DNA replication;
KW Complete proteome.
SQ SEQUENCE 1115 AA; 125349 MW; E8B04E3398E512FE CRC64;

Query Match 4.0%; Score 94; DB 1; Length 1115;

Best Local Similarity 21.7%; Pred. No. 13; Mismatches 155; Indels 144; Gaps 26;

Matches 100; Conservative 62; Mismatches 155; Indels 144; Gaps 26;

20 AIQOAKAVYN-FVAIVDEVREND-----LGQDNELPIDVQASQTASASTDTAN- 66
437 ADQLAKLIPSRGWTLDARQOSPOLDKRLRESSLQOYYSIAKKIEGIPRHASTHAGV 496
67 -----PLDEHEPLLYTT--ALENKTMLINCALNODIMRLACVDTLVHETPAV 113
497 VLSEEPIDVYVLOEGHEBIVLYQYAMDH---LEDIGLKMFLGLRLN-LLIESTT-SM 551
114 IKTKRSIRLDEFTIQTIKGPQVYVQETDPIFLMGNEKGMLTRKDAK---OLEYAA--- 167
552 IEKEWIKIDLS-----SISYSD--DKTF-----SLSKGDTGTFQLESAGMRS 594
168 ---KQFTPLSLSPDLDNRNTPLMSSRPHNPMVYLPIFMHCKPKRSP--NTPSHARQFTPE 224
595 VLKRLKPSGLE-DIVAUNA---LYRP-GEMENIPLEIDRKHGRARVHPHEDLRILEDT 649
225 F-----RAPELKFOYSVYVKAEDMAGTDSLDFGTYOOSHMO 262
650 YGVITYQEDIMIASMAGFSLGEADLLRAVSKKKELLD-----RERSHP- 696
263 IFNKNRSPFRVHDYOPEIFLTPYVSDLPMDGKVRMIGMGAHVHNSGBASAKLSRWNR 322
697 -VEGCLAKREYSV-DTANENV-----DLIYKFAVYG-----FNRSNAVA 732
323 YIMAGHEMKNLTVPRIMGRIRKEGSGSQPDNDPILDYGY----GDVRFYQLENKSN 378

DB 733 YSMIGO---LATLKAHYPIYFMCGLLTSTYIGMEDKISQVLYEAKSGCITILPSPVKKSS 789
 QY 379 I-----SGTVRNPNRSGKALQLDVYPLGKIGSYFQIFQ 414
 DB 790 FPFVENSQVSVSLRAIKSV-----GVSAYVDIYK 819
 RESULT 12
 SRK6_BRAOL STANDARD; PRT; 849 AA.
 AC 009092;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE PUTATIVE SERINE/THREONINE KINASE RECEPTOR PRECURSOR (EC 2.7.1.37)
 DE (S-RECEPTOR KINASE) (SRK).
 GN SRK6.
 OS Brassica oleracea (Cauliflower).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Brassica.
 OX NCBI_TaxID=3712;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. S656; TISSUE=stigma;
 RA MEDLINE=92020942; PubMed=161543;
 RX Stein J.C., Howlett B., Boyes D.C., Nasrallah M.E.;
 RT "Molecular cloning of a putative receptor protein kinase gene encoded
 at the self-incompatibility locus of Brassica oleracea.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:8816-8820(1991).
 CC -!- FUNCTION: INVOLVED IN SPOROPHYTIC SELF-INCOMPATIBILITY SYSTEM
 (THE INABILITY OF FLOWERING PLANTS TO ACHIEVE SELF-
 FERTILIZATION), PROBABLY ACTING IN COMBINATION WITH S-LOCUS-
 SPECIFIC GLYCOPROTEINS. INTERACTION WITH A LIGAND IN THE
 EXTRACELLULAR DOMAIN TRIGGERS THE PROTEIN KINASE ACTIVITY OF THE
 CYTOPLASMIC DOMAIN.
 CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN - ADP + A PHOSPHOPROTEIN.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- TISSUE SPECIFICITY: PREDOMINANTLY IN THE PISTIL AND ANTHER.
 CC -!- POLYMORPHISM: THERE ARE A NUMBER OF DIFFERENT S ALLELES IN
 B. OLERACEA, POSSIBLY PROVIDING THE RECOGNITION SPECIFICITY.
 CC -!- SIMILARITY: THE EXTRACELLULAR DOMAIN IS SIMILAR TO S-LOCUS
 GLYCOPROTEINS OF BRASSICA, WHILE THE INTRACELLULAR DOMAIN IS
 A SER/THR-PROTEIN KINASE RELATED TO RAF KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M76647; AAA33000.1; ALT_TERM.
 DR HSSP: P11362; IEG1.
 DR InterPro: IPR001480; B_lectin.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR003609; Pan_app.
 DR InterPro: IPR002290; Ser_thr_kin_actsite.
 DR InterPro: IPR000858; Slocus_glycop.
 DR Pfam: PR00069; pkinase; 1.
 DR Pfam: PR00954; S_locus_glycop; 1.
 DR SMART: SM00108; B_lectin; 1.
 DR SMART: SM00473; PAN_AP; 1.
 DR SMART: SM00221; STYKC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_SF; 1.
 DR PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
 DR Transferase: Serine/threonine-protein kinase; Signal: ATP-binding;
 KW Transmembrane; Receptor; Glycoprotein; Self-incompatibility.
 FT SIGNAL 1 32
 FT CHAIN 33 849 PUTATIVE SERINE/THREONINE KINASE

FT DOMAIN 33 446 RECEPTOR.
 FT TRANSMEM 447 466 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 467 849 POTENTIAL.
 FT DOMAIN 528 779 CYTOPLASMIC (POTENTIAL).
 FT NP_BIND 534 542 PROTEIN KINASE.
 FT BINDING 556 556 ATP (BY SIMILARITY).
 FT ACT_SITE 653 653 ATP (BY SIMILARITY).
 FT CARBOHYD 47 47 BY SIMILARITY.
 FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 389 389 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 849 AA; 97231 MW; 7E156059EDDF4370 CRC64;
 Query Match 3.9%; Score 92.5; DB 1; Length 849;
 Best Local Similarity 21.9%; Pred. No. 12; Mismatches 121; Gaps 24;
 Matches 78; Conservative 50; Mismatches 107; Indels 121; Gaps 24;
 QY 105 LVHGETPAVVKTKRSIRLDEITIMQTIKQPVVYQETTPDIFLNG---NEK---GMLTK 157
 DB 24 LIHPALSTIYINT-----LSTFESTLISSNKTLV---SPGSFEVGFPRFTRNRMVGLGMWYK 75
 QY 158 K-DKQLEVAKQFTPLSLFSD-----LDRNNTPLMSSRPINPMVLPDIFMNGK 205
 DB 76 KVSDFRTVWVANRPNPLSNAIGLKISGNMLVLDHNSKRPWMWNN-----LTRGN 125
 QY 206 PNRSP-----NTPSHARQFT-----NNEPRAPLKFQVSKVKAADLW 245
 DB 126 -ERSPVVAELLANGNEVYMRSSNDASELYMQSFDPFDLPPLMKLGYNLKT----- 177
 QY 246 GTDSDLMEFGYTOOSHMQIFNGKNSRPFVHDYQ-----PEIFLTO---PYUSDLPMDGK 296
 DB 178 GLNRFL-----TSWRSSDDPSGNNF---SKLETPQSLPEFYLSRENFPNHRSGPNNG- 226
 QY 297 VRMIGKGAHV-----HSNGESAKLSKSNKRAY-----LMAQNEKMLTYMP--RI 339
 DB 227 IREFGIDEDOKLSYVYNEFIENNEVAYTFEMTNSEYSRLLTLLSEGYFORLTWYPSIRI 286
 QY 340 WGRIFKFGSGSQPDNDNDIIDY-----YGVGVAFVLEKENSKNSIGTVR-YNPRS 389
 DB 287 WNRWF-----SSPYD-PQCDTYIMCGPRAYCDV-----NTSPVCNCIGQFNFRN 329
 RESULT 13
 MET_MOUSE
 ID MET_MOUSE STANDARD; PRT; 1379 AA.
 AC P16056; O62125;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HEPATOCYTE GROWTH FACTOR RECEPTOR PRECURSOR (EC 2.7.1.112) (MET PROTO-
 DE ONCOGENE TYROSINE KINASE) (C-MET) (HGF RECEPTOR) (HGF-SF RECEPTOR).
 GN MET.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88262253; PubMed=2838789;
 RA Chan A.M.L., King H.W.S., Deakin E.A., Tempest P.R., Hilkens J.,
 RA Kroezen V., Edwards D.R., Wills A.J., Brookes P., Cooper C.S.;
 RT "Characterisation of the mouse met proto-oncogene.";
 RL Oncogene 2:593-599(1988).
 [2]
 RP SEQUENCE OF 1199-1270 FROM N.A.
 RX MEDLINE=90152381; PubMed=2482828;
 RA Wilks A.F., Kuban R.R., Hovens C.M., Ralph S.J.;
 RT "The application of the polymerase chain reaction to cloning members

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:18:34 ; Search time 44.45 Seconds
(without alignments)
757.461 Million cell updates/sec

Title: US-09-787-083-6

Perfect score: 2360

Sequence: 1 MKVSLSTLTLSTLSCFALLA.....YHNEATSPGVGLMDNMGL 442

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR 68:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	815	34.5	382	2 E81195	phospholipase A1,
2	815	34.5	409	2 H81831	probable phospholip
3	373	15.8	286	2 B36971	outer membrane pho
4	366	15.5	329	2 D81279	phospholipase A1 (
5	358	15.2	289	2 A36971	outer membrane pho
6	357	15.1	289	1 PSECA1	phospholipase A1 (
7	357	15.1	289	2 E86069	outer membrane pho
8	321.5	13.6	289	2 C36971	outer membrane pho
9	246.5	10.4	355	2 C64582	phospholipase A1 P
10	240.5	10.2	355	2 H71930	probable phospholi
11	110.5	4.7	602	1 TVFTRR	protein kinase (EC
12	109.5	4.6	800	2 A29003	cellulase (EC 3.2.
13	109.5	4.6	822	2 JT0611	cellulase (EC 3.2.
14	108	4.6	824	3 JC7532	cellulase (EC 3.2.
15	105.5	4.5	901	2 T20122	hypothetical prote
16	103	4.4	783	2 JC5467	cellulase (EC 3.2.
17	101	4.3	719	2 A42893	penicillin-binding
18	101	4.3	4307	2 T20721	hypothetical prote
19	99.5	4.2	1658	2 T42642	phosphotriostide 3
20	99	4.2	719	2 S28031	penicillin-binding
21	99	4.2	5005	2 F82884	hypothetical prote
22	99	4.2	5005	2 F82884	hypothetical prote
23	98.5	4.2	660	2 E83656	methionyl-tRNA syn
24	98.5	4.2	810	2 A49744	AMP deaminase (EC
25	98	4.2	719	2 S28034	penicillin-binding
26	98	4.2	719	2 S28032	penicillin-binding
27	97.5	4.1	1478	2 S20117	protein kinase BCK
28	96.5	4.1	797	2 T46737	X-Pro dipeptidyl-P
29	96.5	4.1	838	2 A96557	probable receptor

30	96	4.1	608	2 S28036	penicillin-binding
31	95	4.0	719	2 S28035	penicillin-binding
32	95	4.0	719	2 S28038	penicillin-binding
33	95	4.0	719	2 S28037	penicillin-binding
34	95	4.0	765	2 T35719	chitinase - Strept
35	94.5	4.0	397	2 A35136	cellulase (EC 3.2.
36	94.5	4.0	601	2 T26062	hypothetical prote
37	94.5	4.0	655	2 T26061	hypothetical prote
38	94	4.0	1115	2 D69617	DNA polymerase III
39	93.5	4.0	564	2 T40777	lactic reductase t
40	93	3.9	791	2 H96839	hypothetical prote
41	92.5	3.9	857	1 A41369	S-receptor kinase
42	92.5	3.9	1379	1 S01254	hepatocyte growth
43	92	3.9	324	2 B69521	hypothetical prote
44	92	3.9	888	2 S50801	AMP deaminase homo
45	91.5	3.9	486	2 S30959	gene 14 protein -

ALIGNMENTS

RESULT 1
E81195
phospholipase A1, probable NMB0464 [imported] - Neisseria meningitidis (strain MC58 s
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: E81195
R:Telletlin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
Hickey, E.R.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
rl, H.; Qin, H.; Yamathayan, J.; Gill, J.; Scariato, V.; Maignanl, V.; Pizze, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755
A:Accession: E81195
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-382 <TEP>
A:Cross-references: GB:AE02403; GB:AE02098; NID:g7225688; PID:NMF40901.1; PID:g722
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0464

Query Match	34.5%	Score 815	DB 2	Length 382
Best Local Similarity	44.7%	Pred. NO. 1.9e-58		
Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;				
OY 87	INCSALNDIMRIACYDVLHGEPPAVT-----KTKRSIRLDETITWOTI-RGKPOVYQOE 140			
DB 33	LOCALALDNTVRLACYDRIFAAQLPSAGQESKAVNLTEVRSYLDKGEAVIYVEK 92			
OY 141	TTDFIFLGNKKGMLTKKDAKOLEYAKQTFPLSLSPFLDRNN--TPLWSSRPNNMYLVP 199			
DB 93	GGDAL-----PADSAGETADITVPLSLMYLDKNDKGLGVRHNNMYLMP 139			
OY 200	IFMKGKPNRSPNTPSHEAR-OFTPNFRAPPEIKFOVSVKVAEDLWGTDSDLAFGYTQQ 258			
DB 140	LMYNNSPNYAPAGSPRTGTVOEKFGQKRAETKISAEKIAEDLFTFRADLWGYTOR 199			
OY 259	SHWOTFN-GKNSRPFRVADYDPEITFLTOPVYSDLPMDCKVMIMGAVHNSGESAKLSR 317			
DB 200	SDMOIYNGRKSAPPRNDYRPEITFLTOPVAKADLPFGGLRLMAGFVHQSGQSRPSR 259			
OY 318	SMNRAYLMAAGMEKMLTVMPRIWGRIFKESGSGQPDNDLIDYGYGDVFLYOLENKS 377			
DB 260	SMNRITMAAGMEKMLTVIPVWVWRAFDQ-SGDK-NMDPRADLVGYGDVFLYOLENRO 317			
OY 378	NISGTVRYNPRSGKALQDLVYPLGKISGYFOIPOGYGSLIDYHNEATSPGVGLMN 437			
DB 318	NVSYLRYNPRPTGYGALEAAATFPKGLKGVGRFHGYGESLIDYHNEATSPGVGLMN 377			
OY 438	DWMLG 442			

Db 378 DLDGI 382

RESULT 2
H81831

Probable phospholipase A2 [imported] - Neisseria meningitidis (strain 22491 serogroup H81831)
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: H81831
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, N.; Holtrop, S.; Jags, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, N.; Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491.
A:Reference number: A81775; MUID:20222556
A:Accession: H81831
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-409 <PAR>
A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85240.1; PID:g7380665
A:Experimental source: serogroup A, strain 22491
C:Genetics:
A:Gene: NMA2021

Query Match 34.5%; Score 815; DB 2; Length 409;
Best Local Similarity 44.7%; Pred. No. 2.1e-58;
Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;

QY 87 INCSALNODIMRLACVDTLVHGETPAVI-----KTKRSIRLDETIGQTI-KGKPOVYVCE 140
Db 60 LQCAALTDVNTACTDRIFPAQOLPSSACQEGESKAVLLETYSLSLDKGAAYIVKEK 119
QY 141 TDDPFLFMCNEKGMKTKDAKOLEYAAKQFTPLSFDLDRNN-TPLMSSRPNNPVLP 199
Db 120 GGDAL-----PADSAGETADYITPLSLMYDLDKNDLRGLGVRHNPTLMP 166
QY 200 TTMHGKPNNSPTPSHEAR-QFTPNFRAPELKEQVSVKKAEDIMGDSIDMFGYTOQ 258
Db 167 LWNNSPNYAPGSPTRGTVOEKEGQOKRAETLQVSEFKSIEDLEFKTRADLMFGYTOR 226
QY 259 SHMOJFN-GKNSRPRVHGYOPEIFLTOPYVSLPMDGVRMIGMAVHNHNSGESAKLSR 317
Db 227 SMOQITNQRKSAFNNNTYKPEIFLTOPYKALPFGGRMRMGAGFVHQSNOCSRPESR 286
QY 318 SNNRAYLACMEKMLTVMPRIWGRIFKEGSGSQPDNDPILDYGYGVDFRFLYQLENSK 377
Db 287 SNNRIYAMAGMEKGLTYIPRVVRAFDQ-SGDK-NDNPDIDYMGYGVDFRFLYQLENSK 344
QY 378 NISGTVRYNPRSGKALQLDYVYPLGKISGYFOIGYQOSLIDYVNEATSFVGLMLN 437
Db 345 NIVSYLRNPKTCGYAIEAATYFPIKGLKGVVRFHGYGESLIDYVNHKQIGIGIMFN 404
QY 438 DMWGL 442
Db 405 DLDGI 409

RESULT 3
B36971
Outer membrane phospholipase A (EC 3.1.1.1) precursor - Klebsiella pneumoniae
C:Species: Klebsiella pneumoniae
C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 18-Jun-1999
C:Accession: B36971; S40129
R:Brok, R.G.P.M.; Brinkman, E.; van Bortel, R.; Bekkers, A.C.A.P.A.; Verheij, H.M.; Tomm, J. Bacteriol. 176, 861-870, 1994
A:Title: Molecular characterization of enterobacterial plid genes encoding outer membrane phospholipase A
A:Reference number: A86971; MUID:94131966
A:Accession: B36971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-286 <BRO>
A:Cross-references: EMBL:X76901; NID:g436880; PIDN:CAA54223.1; PID:g436881

A:Note: authors translated the codon AAG for residue 112 as Arg
C:Genetics:
A:Gene: plid
C:Superfamily: bacterial phospholipase A1
C:Keywords: carboxylic ester hydrolase

Query Match 15.8%; Score 373; DB 2; Length 286;
Best Local Similarity 42.1%; Pred. No. 9.6e-23;
Matches 90; Conservative 27; Mismatches 87; Indels 10; Gaps 5;

QY 226 RAPELKVSVKKAADLMGDSIDMFGYTOQSHMOIFNGKSRPRRVNDYQPEIFLTQ 285
Db 80 RKDEVRKQLSLAPLWRLGIDNSLGSASTQKSMQSLNSKESAPREFRYEPQLELGF 139
QY 286 PV-YSDLPMDGKVRMIGMAVHNHNSGESATLSNNRAYLACMEKMLTVMPRIWGRIF 344
Db 140 ATDYQFAGW--TLRDIEMGYNHDSNGRSDPTSSWNRILYRLAONGNMLYEVKPM---Y 194
QY 345 KEGSGSQPDNDPILDYGYGVDFRFLYQLENSKISGTVRYNPRSGKALQLDYVYPLGK 404
Db 195 VVGS---TDDNPDITKYMGYRLKGVYQI-GEAIISSAQGYNNWNTGYGALGVSPITK 250
QY 405 GISGYFOIGYQOSLIDYVNEATSFVGLMLN 438
Db 251 HVRATYQIYSYGESLIDYNFNOTRVGVGLMLN 284

RESULT 4
D81279
Phospholipase A1 (EC 3.1.1.32) Cj1351 [imported] - Campylobacter jejuni (strain NCTC C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 18-Aug-2000
C:Accession: D81279
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chli, N.; Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals
A:Reference number: A81250; MUID:20150912
A:Accession: D81279
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-329 <PAR>
A:Cross-references: GB:AL139078; GB:AL111168; NID:g6968723; PIDN:CAB73778.1; PID:g6966
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: plidA; Cj1351
C:Keywords: carboxylic ester hydrolase

Query Match 15.5%; Score 366; DB 2; Length 329;
Best Local Similarity 33.2%; Pred. No. 4.4e-22;
Matches 105; Conservative 48; Mismatches 125; Indels 38; Gaps 13;

QY 130 IKGKPOVYVQETDPIFLMNEKGMKTKDAKOLEYAAKQFTPLSFDLDRNNT--PLW 187
Db 43 LKNSVLSISOENNSSQATOTONSITIKKEKQ-----DFSRLANLYGENESEFNPLG 96
QY 188 SSRPNMNYVLPPIFMHCKRPSRPTPSHEARQFTPNFRAPELKEQVSVKKAEDLMGT 247
Db 97 IS-SYKNNIFLP-FAVSFNSLVGNKKNSA-----KFQSLVKKRLEENLGL 141
QY 248 DSDLMFGYTOQSHMOIFNGKSRPRRVNDYQPEIFLTOPYV-SDLPMDGKVRMIGMAVH 306
Db 142 DEKYIYATQTSWMOIY--BHSSEFRNTQYQPEIFLDPILYLDYEFNNLR--VGLH 196
QY 307 HSNGESAK--LSSNWAYLACMEKMLTVMPRIWGRIFKEGSGSQPDNDPILDYGY 364
Db 197 ESNKGCDENQSSRWNRIVYSTALYKFLFVRLMYRI--DENKKDDONPAIILHMGN 253
QY 365 GDYRFLYQLENSKISGTVRYNPR--SGKALQLDYVYPLGKISGYFOIGYQOSLI 421
Db 254 FDNVLAY-LGDDYFINDMLRNKLFHNNKGAIOVDLGYDIFNNGIYWLQYFNGYGESLI 312

Query Match 10.2%; Score 240.5; DB 2; Length 355;
Best Local Similarity 25.1%; Pred. No. 7.6e-12;
Matches 83; Conservative 44; Mismatches 109; Indels 95; Gaps 13;

QY 157 KDAKOLEYAKOTPLSLSPDLDRNNTPLMSSRPHNPVLPFLFMHGNKRNPSPTPSHE 216
DB 69 KKYLNMDYLGTFLEFPYHSF-----TFIPQYHNNIN----- 102
QY 217 ARQFTPEEFAPRLKFOVSKYKAAEDLMDGTSDSLMEGYQOOSHWOJFNKNSRPFVHD 276
DB 103 ---YORNEF-----KFOISRVPVFRHIIMTKTLTYQTWFOIYNDPOSAFMBMIN 154
QY 277 YQPEIFLQPVYSDLPMDKGV---RMIGCAVHHSNG-ESAKLSRSMNR----- 321
DB 155 FMEPLLIVYPI-NFKPRGKIGNFSSEIWMOHISNGVGAOCYOPNKEGNPENQPPQ 213
QY 322 -----ALMAG-----MEWK--NLTPMPRIGRIFKEGSGSQPDD- 354
DB 214 PVIYKDVNGCKDVRWGGCRSVSAGNALCFVLVMEKGLKTMVAWYPIV-----PYDQ 265
QY 355 -NPDLIDYGYGVRELY-----OLENKSNIIGTVRYNPRSGKALQLDYVYPLGKI 406
DB 266 SNQLIDYMGYGNAKIDYRGRHHEFLQDYFYQWRYD--RWHGAFRLGYTYRINPFV 323
QY 407 SGYFOJFGYGSGLDYNHHEATSPGVGLMN 437
DB 324 GYIOWPENGIGDGLYEDVFSNRIGGIRLN 354

RESULT 11
TVTRRR
protein kinase (EC 2.7.1.37) raf - rat
N:Alternate names: kinase-related transforming protein raf; raf proto-oncogene protein-s
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 11-Jun-1999
C:Accession: B26126
R:Ishikawa, F.; Takaku, F.; Nagao, M.; Sugimura, T.
Mol. Cell. Biol. 7, 1226-1232, 1987
A:Title: Rat c-raf oncogene activation by a rearrangement that produces a fused protein.
A:Reference number: A26126; MUID:87172791
A:Accession: B26126
A:Molecule type: mRNA
A:Residues: 1-602 <RSH>
A:Cross-references: GB:M15428; NID:g206546; PIDN:AAA42002.1; PID:g206547
C:Genetics:
A:Gene: raf
C:Superfamily: rat protein kinase raf; protein kinase homology
C:Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; serine/threonin
F:301-567/Domain: protein kinase homology <KIN>
F:309-317/Region: protein kinase ATP-binding motif
F:329/Active site: Lys #status predicted
F:453/Binding site: phosphate (Ser) (covalent) (by autophosphorylation) #status predicted

Query Match 4.7%; Score 110.5; DB 1; Length 602;
Best Local Similarity 19.5%; Pred. No. 0.6; Indels 145; Gaps 23;
Matches 93; Conservative 68; Mismatches 171;

QY 29 NPVAFVDEVSENDLGDONELPIDVOSASTDTANPLDEHEPELYTALENKTMILN 88
DB 87 NQKVVLTKKEKLETAQDRNLGI-----QSQFTBAKEELAEKRDILRT---NERLSQE 137
QY 89 CSALNDIMRLACYDVLVHGETPAVYIKTRSIKRLDETIWO--TIKGPQVYQDE----- 140
DB 138 VEYLTEDEVKRL--NEKLESNT---TKGELQKLDELQASDVYKVKYREKRLDEQKELLHN 192
QY 141 -----TDDPIFLMNEKG-----MLTPKDAKOLEYAKOTPLSLSPDLD 180
DB 193 QNSWMLTELTKTDELALALREKGNELTELKCTLENKKEEDALIRSHSESAPSLSS-SSP 251
QY 181 RNNTPLMSSRPHNPVLPFLFMHGNKRNPSPTPSHEARQFTPEEFAPRLKFOVSKYKA 240
DB 252 NNLSPTGWSQPKTP-----VPAQRERARAGSGTGQENKKTIRPGQORDSYYWEIASEVM 304

QY 241 AEDLMDGTSDSLMEGYTOOSH-----QIFNGKSRPFVHDYQPEI----- 281
DB 305 LSTRIGSGS---FGTVYKGMHGDVAVKILKVDPTPEQLQAFNEVAALRKTRHNAIL 361
QY 282 -----FLTQ-----PVYSDL-PMDGKVMY-----GMGAVHHSNESA 313
DB 362 FMCYMTDNLAIYQNEGSSLYKHLVQETKFCMPQLIDAROTAGOMDYLAHKNIIHR 421
QY 314 KLSRSMRAYLMAEMEK---NLTPMPRIGRIFKEGSGSQ----- 351
DB 422 DMKS--NNIFLHEGLGVKIGDFLATVKSrw-----SGSQYEQPTGSVLMNAPEYIR 472
QY 352 -PDDNPOL--DYGYGVDFRFLYOLENKSNIIGTVRYNPRSGKALQLDYVYPLGK 405
DB 473 MODNNPFSFOSDYVSYGIV-LVEL-----MTGELPYSHINRDOI---IFWVGRG 518

RESULT 12
A29003
cellulase (EC 3.2.1.4), alkaline - Bacillus sp.
N:Alternate names: endo-1,4-beta-glucanase
C:Species: Bacillus sp.
C>Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 21-Jul-2000
C:Accession: A29003
R:Fukunori, F.; Kudo, T.; Narahashi, Y.; Horikoshi, K.
J. Gen. Microbiol. 132, 2329-2335, 1986
A:Title: Molecular cloning and nucleotide sequence of the alkaline cellulase gene fro
A:Reference number: A29003; MUID:87085443
A:Accession: A29003
A:Molecule type: DNA
A:Residues: 1-800 <FUK>
A:Cross-references: GB:000066; GB:N00066; NID:g216223; PIDN:BAAD0045.1; PID:g216224
A:Experimental source: strain 1139
C:Function:
A:Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as
C:Superfamily: Bacillus sp. alkaline cellulase; Thermotoga xylanase A amino-terminal
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
F:585-726/Domain: Thermotoga xylanase A amino-terminal repeat homology <TXA>

Query Match 4.6%; Score 109.5; DB 2; Length 800;
Best Local Similarity 19.8%; Pred. No. 1.1; Indels 169; Gaps 25;
Matches 95; Conservative 54; Mismatches 163;

QY 9 TLISLSCFALIAIOAKAVNPVAFVDEVSEND---LGDNELPIDVOSASTDTA 65
DB 7 TKQILSSILILVL-LLSLFPALAAEGNTRDNKKHLGNDNVKRRPEAGALQLEVDGQ 65
QY 66 NPL-DEHEPELY-----TALENKTMILNCSAL-----NODIMRLACY-DVLVHGETP 111
DB 66 MTLVDQGEKIQLRGMSSTHGLQMPPEILNDNAVYKALANDWESNMIRLAMYGENGYASNP 125
QY 112 AVITKRSIRLDETI-----WQT-----IKGKPOVY 138
DB 126 ELKSRVYKIGDIALIENDMYIVDMVHAPRDPDYAGAEFFRIDIALYPPNPHITV 185
QY 139 QETDTP-----FLKNGNEKMLTKKDAKOLEYAAQFTPLSLSPDLDRN---NTPLM 187
DB 186 ELANEPSSNNNGAGIRNNEGMNAVK-----EVADPIVELKRSNGAADNIIIVGSPNW 240
QY 188 SSRP-----HNPMYVLPFL--MHGKPNRS--PNTDSHEARQFTPEEFAPRLKFO 233
DB 241 SQRDLADNPIDIDHHTMYVHFYTGSHAASTESYPRPTEPSEKGNVMSNTRYA-----LE 296
QY 234 VSVYKAAEDLMGT-----SDSLMFGYTOO-----SHWQJFNKNS-----RP 271
DB 297 NGVAVFATE--WGTQANGDGGRPYDEADWIEFLANENNISWAMWSLTLN-KNEVSGAFTP 353
QY 272 FRVH-----DYQPEIFLQPVYSDLPW---D 294
DB 354 FELCKSNATSLDPPDOVWVPEELSLSGEYVRAIRIKGVNIET---IDRTKTYKVLMPND 410

```
QY 295 GKVNMIGCAVHNSNGS-----AKLSRNRAYLMAGMEKNTLMPRING 341
DB 411 GTRKQGFV-----NGDSPEVDVVIENEGALKLSGLDSNDVSEGNVAMNRLSADGNG 464
QY 342 R 342
DB 465 K 465

RESULT 13
J0611
cellulase (EC 3.2.1.4), alkaline - Bacillus sp. (strain KSM-64)
N:Alternate names: endo-1,4-beta-glucanase
C:Species: Bacillus sp.
C:Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 10-Dec-1999
C:Accession: J0611
R:Sumitomo, N.; Ozaki, K.; Kawai, S.; Ito, S.
Biosci. Biotechnol. Biochem. 56, 872-877, 1992
A:Title: Nucleotide sequence of the gene for an alkaline endoglucanase from an alkalophilic
A:Reference number: J0611; MUID:92305459
A:Accession: J0611
A:Molecule type: DNA
A:Residues: 1-822 <SUM>
A:Cross-references: GB:M64963; MID:g289264; PIDN:AAV3189.1; PID:g289266
C:Function:
A:Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as cel-
A:Pathway: cellulose degradation
C:Superfamily: Bacillus sp. alkaline cellulase; Thermotoga xylanase A amino-terminal ref
C:Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation
F:583-726/Domain: Thermotoga xylanase A amino-terminal repeat homology <TXA>

Query Match 4.6%; Score 109.5; DB 2; Length 822;
Best Local Similarity 19.8%; Pred. No.1.2; Mismatches 163; Indels 169; Gaps 25;
Matches 95; Conservative 54;

QY 9 TSLISCFALIAIOQAKAVNPVAFVDEVSEND---LCQDNELPIDVQATQASASTDA 65
DB 7 TKQLISSILILVL-LTSLPPALAEAGNTRDNFKHLGNDNVKRPSEAGALQLOEVQDQ 65
QY 66 NPL-DEHEPELY-----TLALENKTMLINCAL-----NODIMLACY-DLVHGEIP 111
DB 66 MTLVLDQHGKIDLRGKSTHGLQWPEPLNDNAVKALANDMESMIRLAMYGEENGASNP 125
QY 112 AVIKTKRSIRLDETI-----WQT-----IKGKPOVY 138
DB 126 ELIKSRVIGIDLALENDMYVIVDMVHVARGPDPVYAGADEFFRDIAALYPPNPHIT 185
QY 139 QETTPD-----IFLMGNEKGLTKKQAKOLEYAKOFTPLSLFDDLRN---NTPLM 187
DB 186 ELANPESSNNNGAGIPNNEEGMNAVK-----EYADPIVEMLRDSGNADNDNIIIVGSPWM 240
QY 188 SSRP-----HNPMVVLPIF--MHGKPNRS--PNTPSHEARQFTNEFRAPBLKQ 233
DB 241 SCRPDLADNPIDDHHTMTVHFYTGSHAASTESTPEPNSRGVMSNTRYA---LE 296
QY 234 VSVKVAKAEDLMGT-----DSDLWFGYTOQ-----SHMOIFNGKNS---RP 271
DB 297 NCVAVAFATE--WGTSQANGDGGPYDEADVWIEFLNENNISMAMWSLTN-KNEVSATFP 353
QY 272 FNVH-----DYQPIFLTPVYSDLPV---D 294
DB 354 FELGKSNATSLDPGPQVAVPEELSLSGEYVARIKGVNYP---IDRTKYTKVLWDFND 410
QY 295 GKVNMIGCAVHNSNGS-----AKLSRNRAYLMAGMEKNTLMPRING 341
DB 411 GTRKQGFV-----NGDSPEVDVVIENEGALKLSGLDSNDVSEGNVAMNRLSADGNG 464
QY 342 R 342
DB 465 K 465
```

```
RESULT 14
J07532
cellulase (EC 3.2.1.4), alkaline - Bacillus sp. (strain KSM-5237)
N:Alternate names: endo-1,4-beta-glucanase
C:Species: Bacillus sp. (strain KSM-5237)
C:Date: 31-Mar-2001 #sequence_revision 31-Mar-2001 #text_change 31-Mar-2001
C:Accession: J07532; PC7107
R:Hakamada, Y.; Hatada, Y.; Koike, K.; Yoshimatsu, T.; Kawai, S.; Kobayashi, T.; Ito,
Biosci. Biotechnol. Biochem. 64, 2281-2289, 2000
A:Title: Deduced amino acid sequence and possible catalytic residues of a thermostabl
A:Reference number: J07532; MUID:21036886
A:Accession: J07532
A:Molecule type: DNA
A:Residues: 1-824 <HAK>
A:Cross-references: DDBJ:AB018420
A:Experimental source: strain KSM-5237
A:Accession: PC7107
A:Molecule type: protein
A:Residues: 31-50 <HA2>
C:Comment: This enzyme is a thermostable, alkaline enzyme used as effective detergent
C:Genetics:
A:Gene: Egl-237
C:Keywords: hydrolase; glycosidase

Query Match 4.6%; Score 108; DB 3; Length 824;
Best Local Similarity 19.8%; Pred. No.1.5; Mismatches 159; Indels 166; Gaps 28;
Matches 95; Conservative 60;

QY 9 TSLISCFALIAIOQAKAVNPVAFVDEVSEND---LCQDNELPIDVQATQASASTDA 65
DB 7 TKQLISSILILVL-LTSLPPALAEAGNTRDNFKHLGNDNVKRPSEAGALQLOEVQDQ 65
QY 66 NPL-DEHEPELY-----TLALENKTMLINCAL-----NODIMLACYDLVHGE--- 109
DB 66 MTLVLDQHGKIDLRGKSTHGLQWPEPLNDNAVKALSNDSMIRLAMY---VGENCY 121
QY 110 ---TPAVIKTK-----RSIRLDETI---WQT-----IKGK 134
DB 122 ATNPBLIKORVIGIDLELALENDMYVIVDMVHVARGPDPVYAGAKDFPRETALYPPNP 181
QY 135 QVYVOETTPD-----IFLMGNEKGLTKKQAKOLEYAKOFTPLSLFDDLRN---N 183
DB 182 HIIVELANPESSNNNGAGIPNNEEGMNAVK-----EYADPIVEMLRKSGNADNDNIIIVG 236
QY 184 TPLWSSRP-----HNPMVVLPIFEMHGKPNRS--PNTPSHEARQFTNEFRAPBLK-- 231
DB 237 SPWMSQRPDLADNPIDDHHTMTVHFYTGSHAASTESTPSE---TPNSEKGVMSNT 291
QY 232 ---FQVSVKVAKAEDLMGT-----DSDLWFGYTOQ-----SHMOIFNGKNS-- 269
DB 292 RYALENGVAVFATE--WGTSQASGCGPYFEADVWIEFLNENNISMAMWSLTN-KNEVS 348
QY 270 ---RPRPV-----HDYQP-EIPLT-----QPV---YSDLPV--- 293
DB 349 GAFTPELKSNAITNDPPDPDHVMAPEELSLSGEYVARIKGVNYPIDRTKYTKVLWDF 408
QY 294 -DGKVMIGCAVHNSNGSAKLSRNRAYLMAGME-----MKNLVMPRIWIR 342
DB 409 NQGTKO--GFGVNSDSPNKEILAVDNEENTLKVSGLDVSDVSDGNFWMANRLSANGWK 466

RESULT 15
T20122
hypothetical protein F2583.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T20122; T21324
R:McMurray, A.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z19225
A:Accession: T20122
```

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-901 <W11>
A:Cross-references: EMBL:270750; PIDN:CAA94750.1; GSPDB:GN00023; CESP:F25B3.1
A:Experimental source: clone C50F4
R:Gardner, A.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z19406
A:Accession: T21324
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-901 <W12>
A:Cross-references: EMBL:270752; PIDN:CAA94758.1; GSPDB:GN00023; CESP:F25B3.1
A:Experimental source: clone F25B3
C:Genetics:
A:Gene: CESP:F25B3.1
A:Map position: 5
A:Introns: 25/3; 47/2; 199/1; 242/1; 315/2; 544/3; 600/1; 626/1; 646/1; 666/3; 706/3; 75

Query Match 4.5%; Score 105.5; DB 2; Length 901;
Best Local Similarity 19.6%; Pred. No. 2.8;
Matches 74; Conservative 60; Mismatches 174; Indels 69; Gaps 15;

```
QY 31 VAFVDEVRSENDLGO-----DNELPIDVQSATOSASTDTANPLIDEHEPELYTTALENK 83
DB 468 VAMITEIRNQKLEAEVDRHPIPEVPTPQLASRNPALNOPTDADDDAETSNMRFERS 527
QY 84 TMLI-----NCSALNODIMRLACYDTLVHGEPPIAVIKTKRSIRLDETIMQT 129
DB 528 NVSITMVTGVAIRASNSASPSSKRDQLQRARDLIEKSTTPAATPNSRKASDEERRREE 587
QY 130 I-----KGPQVYVYQETTDPIFLMGNEKMLTKKDAKQLEYAAKQF-----TPLSLSPD 178
DB 588 VRRMLNEKHOPTAIPSTSSSPYPTERRIDGSNT--DLRRIELDVHKKKRDPSPTLVRRQ 645
QY 179 LDRNNTPLMSSRPHNMYLPIFMHCKPNRSPV--TPSHEA--RQFTPNEFRAPELKF 232
DB 646 YDPNDI-----PH-----VPAIGRGRTNGNRNDSSTPSSASTFEDRVKRYGSMRSALKE 694
QY 233 QVSVKVK---AAEDLMGTDSLDMFGYTOQ--SHMQ---IFNGKNSRPFRVHDYQPEIF 282
DB 695 SIQIMAKQGYGMGNDESSSQDALATPTKFFSQWMEKDVDDVEGTANLVLRIDERSDIT 754
QY 283 LTPQVYSDLPMQGVKVMIGMAVHSHNGESAKLSRSMNRAYLMAGMEWKN-LTVPPIWG 341
DB 755 AOADVYID-----KIRETEVGSSEEMLTASYLELTNERNTLVHROEYNIETIRQVTS 809
QY 342 RIFKES--GSQPDNDP 356
DB 810 EIDQLGKQINEVPDDFP 826
```

Search completed: November 30, 2001, 14:18:35
Job time: 257 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:17:48 ; Search time 36.79 Seconds
(without alignments)
270.358 Million cell updates/sec

Title: US-09-787-083-8

Perfect score: 2363

Sequence: 1 MKVSLSTLSTLSTLPCFAILA.....YHNHATSPGVGLMNDMGL 442

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCYUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100.5	4.3	1726	2	US-08-609-049A-30
2	100.5	4.3	1726	4	US-09-170-996-30
3	99.5	4.2	1658	2	US-08-609-049A-13
4	99.5	4.2	1658	4	US-09-170-996-13
5	95.5	4.0	682	3	US-08-481-435-6
6	90.5	3.8	857	1	US-07-717-331F-2
7	89.5	3.8	1024	4	US-09-091-117-5
8	88.5	3.7	503	1	US-07-946-497-2
9	88.5	3.7	503	1	US-08-483-332-2
10	88.5	3.7	503	2	US-08-478-882-2
11	87.5	3.7	666	4	US-08-961-083-2
12	86	3.6	781	1	US-08-373-134D-2
13	86	3.6	781	2	US-09-114-637-2
14	86	3.6	985	5	PCT-US96-03916-6
15	86	3.6	985	5	PCT-US96-03916-6
16	85	3.6	816	1	US-07-731-157A-4
17	85	3.6	816	1	US-08-229-444B-2
18	85	3.6	816	2	US-08-541-780-4
19	85	3.6	1627	1	US-07-665-792E-9
20	84.5	3.6	774	3	US-08-902-632-2
21	84.5	3.6	774	3	US-09-073-354-1
22	84.5	3.6	774	3	US-08-656-005A-1
23	84.5	3.6	774	4	US-09-073-259-1
24	84.5	3.6	774	4	US-09-363-095-1
25	84.5	3.6	774	4	US-09-418-027-1
26	84	3.6	522	6	RE34606-6
27	83.5	3.5	657	4	US-09-306-593-2

28	83	3.5	1088	3	US-08-633-768A-1	Sequence 1, Appli
29	82.5	3.5	535	2	US-08-564-972-1	Sequence 1, Appli
30	82.5	3.5	535	4	US-09-171-966-9	Sequence 9, Appli
31	82.5	3.5	797	4	US-09-086-912-2	Sequence 2, Appli
32	82	3.5	320	2	US-08-245-511-4	Sequence 4, Appli
33	82	3.5	320	2	US-08-600-993A-4	Sequence 4, Appli
34	82	3.5	355	1	US-07-946-497-5	Sequence 5, Appli
35	82	3.5	355	1	US-08-483-322-5	Sequence 5, Appli
36	82	3.5	355	2	US-08-478-882-5	Sequence 5, Appli
37	82	3.5	1022	3	US-08-772-270A-2	Sequence 16, Appli
38	81.5	3.4	1381	4	US-09-540-245A-16	Sequence 7, Appli
39	81.5	3.4	663	4	US-08-441-139-7	Sequence 68, Appli
40	81.5	3.4	693	4	US-08-235-836C-68	Sequence 6, Appli
41	81.5	3.4	844	1	US-07-731-157A-6	Sequence 6, Appli
42	81.5	3.4	844	2	US-08-541-780-6	Sequence 6, Appli
43	81	3.4	334	2	US-08-359-850-4	Sequence 4, Appli
44	81	3.4	537	2	US-08-633-879C-2	Sequence 2, Appli
45	81	3.4	649	4	US-09-618-419-22	Sequence 22, Appli

ALIGNMENTS

RESULT 1
US-08-609-049A-30

Sequence 30, Application US/08609049A

Patent No. 5948664

GENERAL INFORMATION:

APPLICANT: Williams, Lewis T.

APPLICANT: Molz, Lisa

TITLE OF INVENTION: NO. 5948664e1 PI 3-Kinase Polypeptides

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/609, 049A

FILING DATE: 29-FEB-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Dow, Karen B.

REGISTRATION NUMBER: 29, 684

REFERENCE/DOCKET NUMBER: 2307K-063700US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-326-2400

TELEFAX: 415-326-2422

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 1726 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-609-049A-30

Query Match 4.3%; Score 100.5; DB 2; Length 1726;

Best Local Similarity 18.9%; Pred. No. 0.62;

Matches 86; Conservative 66; Mismatches 145; Indels 159; Gaps 21;

QY 45 QDNELPIDV-----OSATOSASTDTANPLDE-----HPELEYTALENKMLIN----- 88

DB 563 EDDDEAPVDINKYLYQLEKRYKEMVRHPPVEELLDSDYHYQVEL-ALQTEHQHRAVDQVIKA 621

```

QY 89 -----CSALNODIMRLACDYLIVHGETP-----AVIKTRKSIRLDETIWOTIKGKPOVYYQE 140
    |||:
    |||:
Db 622 VRKICSLD-----GVETPSVTEAVKRLKRAVLP----- 652
QY 141 TTDFIFLWNGEKMGLTKKDAKOLEYAAKQFTPLSLSPDLDRNNPLMSSRHPNMYLP- 199
    |||:
    |||:
Db 653 -----NKSADVT-----SLGSDTRKNTKSGSLNENPVQVSM 686
QY 200 -----IFMGRKPNR-----SPNTPSHEAKQFTPNFRAPRLKFOVSVYKAAEDL- 244
    |||:
    |||:
Db 687 HLTTRIYDLRLHANSRSGTCGPRGSRNKEAWTATE-----QLQFTVYAHGISS 738
QY 245 -WGTDSDLMFGYTOOSHWOIFNGKN-SRPF--RVHDYQPEIFLTQ-----PV-YSDL 291
    |||:
    |||:
Db 739 NWVSNYEKYIYICLSLH-----NGKDLFRPIOSKRVGTGKNEFYLIKWDELIIFFIQLISOL 794
QY 292 PMDGKVRMIGMAVHNSGESAKLSRSWN-----RAYLMAEMKKNLTVM 336
    |||:
    |||:
Db 795 PLESVHLTLFGLVNLQSSGSSPDSSNKQKRGPEALGKVSLLTFDKRFLTCG-----TKL 848
QY 337 PRWGRIFEKSGSGOPDPNDILDYGYGDVRFPLYOLENKNISGTVAYNRSKGALQL 396
    |||:
    |||:
Db 849 LYLW-----TSSHTNSIPALPKKSYMERIVLYQVDFPSPADIIYTSPOIDRNIIQQ 901
QY 397 DYVYPLGKISGYFOIFGYGOSLIDYNHEATSGV 432
    |||:
    |||:
Db 902 DKLETLESIDIK-----KLIDIHRDSSFGL 927

RESULT 2
US-09-170-996-30
; Sequence 30, Application US/09170996
; Patent No. 6291220
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Moliz, Lisa
; APPLICANT: Chen, Yen-Wen
; TITLE OF INVENTION: No. 6291220e1 PI 3-Kinase Polypeptides
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/170,996
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/609,049
; FILING DATE: 29-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-063700US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1726 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-170-996-30

```

```

Query Match 4.3%; Score 100.5; DB 4; Length 1726;
Best Local Similarity 18.9%; Pred. No. 0.62;
Matches 86; Conservative 66; Mismatches 145; Indels 159; Gaps 21;

QY 45 QDNELPIDV-----QSATQASASTDANPLDE-----HEPILYTALENKTMLIN----- 88
    |||:
    |||:
Db 563 EDDAPVDLNLKLYLOIEKPYKEVMTRHPEBELDSYNQVEL-ALQENOHRAVDYIKA 621
QY 89 -----CSALNODIMRLACDYLIVHGETP-----AVIKTRKSIRLDETIWOTIKGKPOVYYQE 140
    |||:
    |||:
Db 622 VRKICSLD-----GVETPSVTEAVKRLKRAVLP----- 652
QY 141 TTDFIFLWNGEKMGLTKKDAKOLEYAAKQFTPLSLSPDLDRNNPLMSSRHPNMYLP- 199
    |||:
    |||:
Db 653 -----NKSADVT-----SLGSDTRKNTKSGSLNENPVQVSM 686
QY 200 -----IFMGRKPNR-----SPNTPSHEAKQFTPNFRAPRLKFOVSVYKAAEDL- 244
    |||:
    |||:
Db 687 HLTTRIYDLRLHANSRSGTCGPRGSRNKEAWTATE-----QLQFTVYAHGISS 738
QY 245 -WGTDSDLMFGYTOOSHWOIFNGKN-SRPF--RVHDYQPEIFLTQ-----PV-YSDL 291
    |||:
    |||:
Db 739 NWVSNYEKYIYICLSLH-----NGKDLFRPIOSKRVGTGKNEFYLIKWDELIIFFIQLISOL 794
QY 292 PMDGKVRMIGMAVHNSGESAKLSRSWN-----RAYLMAEMKKNLTVM 336
    |||:
    |||:
Db 795 PLESVHLTLFGLVNLQSSGSSPDSSNKQKRGPEALGKVSLLTFDKRFLTCG-----TKL 848
QY 337 PRWGRIFEKSGSGOPDPNDILDYGYGDVRFPLYOLENKNISGTVAYNRSKGALQL 396
    |||:
    |||:
Db 849 LYLW-----TSSHTNSIPALPKKSYMERIVLYQVDFPSPADIIYTSPOIDRNIIQQ 901
QY 397 DYVYPLGKISGYFOIFGYGOSLIDYNHEATSGV 432
    |||:
    |||:
Db 902 DKLETLESIDIK-----KLIDIHRDSSFGL 927

RESULT 3
US-08-609-049A-13
; Sequence 13, Application US/08609049A
; Patent No. 5948664
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Moliz, Lisa
; APPLICANT: Chen, Yen-Wen
; TITLE OF INVENTION: No. 5948664e1 PI 3-Kinase Polypeptides
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,049A
; FILING DATE: 29-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-063700US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:

```

LENGTH: 1658 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-609-049A-13

Query Match 4.2%; Score 99.5; DB 2; Length 1658;
Best Local Similarity 18.9%; Pred. No. 0.74;
Matches 101; Conservative 79; Mismatches 194; Indels 159; Gaps 23;

QY 5 LSTLTSLIPCFALIAIQ--QAQAVNPVAFVDEVS-----ENDLGQNE 48
DB 381 LSPVTVQRMNCGENASVKSIEIGLQLPVTFCDVSVTEIIMQALCWVHDLNO--- 437
QY 49 LPIDVOS-----ATQASSTDPANPLDEHE--PELYTTALENKTMLINCSALNDIMRLAC 101
DB 438 --VDVGSYILKVGQGEVYLNHCLGSHHEIQNCRKMDTEIKQLLTLISAMQNLARTAE 495
QY 102 YDPLVHGETPVAVTKRSTRIDETIMQTKGKQVYQETTPIF----- 146
DB 496 DD-----EAP-----VDLNKLYQIEKPYKEVMIRHPEELLDVSYHYQVELALQTEN 542
QY 147 -----LMGNEKGM/LTKDAKOLEYAA-----KQTPPLSLSDLDNRN 183
DB 543 QHRAVDQVIAKVRKICSAIDGVEPVSYTEA-VKIKRAVNLPRNKSADVTSLSGSDTRKN 601
QY 184 TPLMSSRPHPMYVLP-----IFMKGKPNR-----SPNTPSHKAKOPTPNEFRAP 229
DB 602 STKGSILNPEPNVQVSMDLTATAYDLRLHANSSRCSTGCPGRSRNKEAWTATE----- 656
QY 230 LKQVSVKKAABDL--WGTSDLMFQYTOQSHMOIFNGKN-SRPF--RVHDIQPEIF 282
DB 657 --QLOFTVYAAHGISSNMWNSYKYYLCLSLH---NGKDLFKPIQSKKVGYYKNFEY 709
QY 283 LTPQ-----PV-YSDLPMDGKVRMIGMGAHVHNSGESAKLSRSWN----- 320
DB 710 LIKWDELIFPIQISQPLESVLHLTLFGLVNLQSSGSSPDSNKRQKPEALGKVSITLFD 769
QY 321 -RAYLAGMEKMKLTVMPRIMGRIFKESGSGOPDNDIIDDYGYGVDFLYOLENKSNI 379
DB 770 FKRFITCG-----TKLLYLM-----TSSHTNSIPGALPKKSYVERIVLQVDFPSPA 816
QY 380 SGTVRYNPRSGKALQLDVYVPLGKIGSYFOIRQGYGOSLIDYNHETATSGV 432
DB 817 FDIITYSPQIDRNIIQDKLETLESDIK-----KILDIIHSDSSFGL 859

RESULT 4

US-09-170-996-13
Sequence 13, Application US/09170996
Patent No. 6291220
GENERAL INFORMATION:
APPLICANT: Williams, Lewis T.
APPLICANT: Moliz, Lisa
APPLICANT: Chen, Yen-Wen
TITLE OF INVENTION: No. 6291220e1 PI 3-kinase Polypeptides
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/170,996

FILING DATE: 435
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/609,049
FILING DATE: 29-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-06370005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ. ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1658 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-170-996-13

Query Match 4.2%; Score 99.5; DB 4; Length 1658;
Best Local Similarity 18.9%; Pred. No. 0.74;
Matches 101; Conservative 79; Mismatches 194; Indels 159; Gaps 23;

QY 5 LSTLTSLIPCFALIAIQ--QAQAVNPVAFVDEVS-----ENDLGQNE 48
DB 381 LSPVTVQRMNCGENASVKSIEIGLQLPVTFCDVSVTEIIMQALCWVHDLNO--- 437
QY 49 LPIDVOS-----ATQASSTDPANPLDEHE--PELYTTALENKTMLINCSALNDIMRLAC 101
DB 438 --VDVGSYILKVGQGEVYLNHCLGSHHEIQNCRKMDTEIKQLLTLISAMQNLARTAE 495
QY 102 YDPLVHGETPVAVTKRSTRIDETIMQTKGKQVYQETTPIF----- 146
DB 496 DD-----EAP-----VDLNKLYQIEKPYKEVMIRHPEELLDVSYHYQVELALQTEN 542
QY 147 -----LMGNEKGM/LTKDAKOLEYAA-----KQTPPLSLSDLDNRN 183
DB 543 QHRAVDQVIAKVRKICSAIDGVEPVSYTEA-VKIKRAVNLPRNKSADVTSLSGSDTRKN 601
QY 184 TPLMSSRPHPMYVLP-----IFMKGKPNR-----SPNTPSHKAKOPTPNEFRAP 229
DB 602 STKGSILNPEPNVQVSMDLTATAYDLRLHANSSRCSTGCPGRSRNKEAWTATE----- 656
QY 230 LKQVSVKKAABDL--WGTSDLMFQYTOQSHMOIFNGKN-SRPF--RVHDIQPEIF 282
DB 657 --QLOFTVYAAHGISSNMWNSYKYYLCLSLH---NGKDLFKPIQSKKVGYYKNFEY 709
QY 283 LTPQ-----PV-YSDLPMDGKVRMIGMGAHVHNSGESAKLSRSWN----- 320
DB 710 LIKWDELIFPIQISQPLESVLHLTLFGLVNLQSSGSSPDSNKRQKPEALGKVSITLFD 769
QY 321 -RAYLAGMEKMKLTVMPRIMGRIFKESGSGOPDNDIIDDYGYGVDFLYOLENKSNI 379
DB 770 FKRFITCG-----TKLLYLM-----TSSHTNSIPGALPKKSYVERIVLQVDFPSPA 816
QY 380 SGTVRYNPRSGKALQLDVYVPLGKIGSYFOIRQGYGOSLIDYNHETATSGV 432
DB 817 FDIITYSPQIDRNIIQDKLETLESDIK-----KILDIIHSDSSFGL 859

RESULT 5

US-08-481-435-6
Sequence 6, Application US/08481435
Patent No. 6027906
GENERAL INFORMATION:
APPLICANT: Balganes, Tanjore S
APPLICANT: Town, Christine
TITLE OF INVENTION: No. 6027906e1 Polypeptides
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:

ADDRESSER: White & Case
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,435
FILING DATE: 10-JUL-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: IN 580/MAS/94
FILING DATE: 01-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9404072-2
FILING DATE: 24-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Steiner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-151
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 682 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-481-435-6

Query Match 4.08; Score 95.5; DB 3; Length 682;
Best Local Similarity 19.98; Pred. No. 0.45;
Matches 89; Conservative 69; Mismatches 186; Indels 103; Gaps 24;

QY 18 ILATQQAQAVNPVAFVDEVRSENDL-GQDNEPLIDVQASQ-----ASTDFANP 67
DB 10 LVATTSKTIIDNKQIADLGSERVANAANDIPTDYKALVISEDHFFHRCIDITIRI 69
QY 68 LDEHEPELYTALLENKTMILNCSALNODIMRLACYDTLVHGETPAVIKTRSIRLDETIV 127
DB 70 LGARLRNLQSSLSGG-----SALTQOLIKLYEFSTSTSDQITIS-----KRAQEW 115
QY 128 QTIKGFQVYQET-----TDPIFLMGNEKMLTK-----KDAKOLF-----YA 166
DB 116 LAIDLEKATKQEIITLYINKVYMSNGYMGTAQANYGKDLNLSLPOLALLAGMPQA 175
QY 167 AKOTPTLS-LSFDDLRRNTPMLSSRPHPVYVLPPIFMGKPNRSPNPSHEAKOFTPEEF 225
DB 176 PNOQDPTSHPEAADRRRLV--SEMKNOGIT-----SAEQEKAVNPITITGLQSLKAS 229
QY 226 RAPP-----LKFQYS-VKVAEDLMDGTDSDLMFGYTQ--QSH-WQIFNGKNSRPFVHD 276
DB 230 NPYAYMONYKKEVINQVBEETGYNLTGTGMVYTNVQDEAKHLMIDINTEDEYVAPDE 289
QY 277 YQPIELFLOPYSDLPMDGKVRMIGMGAVHNSN-----GSAKLSRSWNRAYILMAGHEW 330
DB 290 LQ-----VASTIVVVS-NGKV-IAQLGARHQSNSVSEGINQAVETNRDM-----GSTM 335
QY 331 KNLV-VMPRMIRIGRIKFEKSSGQPDNDPILDYG-----YGVDFEFLYLEKMS 377
DB 336 KPIIDYPALEYGY-ESTATIVHDEP--YNPPTNTVYVMDRGYFNITLQTLAQOSR 392
QY 378 NISGTVRYN-----PRSGKALDLDY 398
DB 393 NPAVETLNKYGILNRAKTFNLGLGIDY 419

RESULT 6
US-07-717-331F-2
Sequence 2, Application US/07717331F
Patent No. 5484905
GENERAL INFORMATION:
APPLICANT: June Nasrallah, Michael Nasrallah, and Joshua
APPLICANT: Stein
TITLE OF INVENTION: A Receptor Protein Kinase Gene
TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: Yahwak & Associates
STREET: 25 Skytop Drive
CITY: Trumbull
STATE: Connecticut
COUNTRY: USA
ZIP: 06611
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: Macintosh
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/717,331F
FILING DATE: June 19th 1991
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: George M. Yahwak
REGISTRATION NUMBER: 26,824
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203)268-1951
TELEFAX: (203)268-1951
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 857 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-717-331F-2

Query Match 3.88; Score 90.5; DB 1; Length 857;
Best Local Similarity 21.44; Pred. No. 2.2;
Matches 76; Conservative 50; Mismatches 110; Indels 119; Gaps 23;

QY 105 LVHGETPAVIKTRSIIRLDETIVQTIKGPQVYVQETTDPIFLMG---NEK---GMLTK 157
DB 24 LIHPLASIIYINT-----LSTESTLTISSNKTLY---SPGSIPEVGFRTNSRWYLGMYWK 75
QY 158 K-DAKQLEYAAKQFTPLSLFD-----LDRNTPMLSSRPHPVYVLPPIFMHGK 205
DB 76 KVSRTYVYVAVNRDPLNLAIGTLKISGNNLVLLDHSKPKVWMTN-----LTRGN 125
QY 206 PNRSP-----NTPSHEAKOFT-----PNEFRAPELKFOVSVKKAEDLM 245
DB 126 -ERSPVVAELLANGFVWRDSSNDASEYLMQSFYPIIDTLPEKLTGYNLKT----- 177
QY 246 GTDSDLMFGYTQOSHQWQIFNGKNSRPFVHDYQ-----PELFIHQ--PYVSDLPMDGK 296
DB 178 GLNRL-----TSWRSSDDPSSGNF--SYKLETLQSLPEVYLSRENFPMHRSQPMWG- 226
QY 297 VRMIGMGAVH-----HSNGESAKLSRSWNRAY-----LMAKGMKNUITYWP--RI 339
DB 227 IFRSGIPEDOKLSYVYVNFIEENNEVAVTFRTNNSFYSRLTLISEGYFORLTWVPSIRI 286
QY 340 WGRIRKESGSGOPD-----NPDIIDYGYGQVRFYOLEKNSNISGTVR-VNPRS 389
DB 287 WNRWSSPVDRQCDITYIMCGP-----YAYCDV-----NTPSPVNCIOGFNPRN 329

RESULT 7
US-09-091-117-5


```

: Sequence 5, Application US/09091117
: Patent No. 6171589
: GENERAL INFORMATION:
: APPLICANT: The University of Melbourne
: TITLE OF INVENTION: Mycoplasma Recombinant Polypeptides and
: TITLE OF INVENTION: Vaccines
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: GREENLEE, WINNER and SULLIVAN P.C.
: STREET: 5370 Manhattan Circle, Suite 201
: CITY: Boulder
: STATE: Colorado
: COUNTRY: United States of America
: ZIP: 80303
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/091,117
: FILING DATE: 12 JUNE 1998
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/AU96/00803
: FILING DATE: 13-DEC-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: AU PN7127
: FILING DATE: 13-DEC-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: WINNER, Ellen P.
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: +1 303 499 8089
: TELEFAX: +1 303 499 8089
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1024 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: Mycoplasma genitalium
: US-09-091-117-5

Query Match 3.8%; Score 89.5; DB 4; Length 1024;
Best Local Similarity 20.2%; Pred. No. 3.7;
Matches 86; Conservative 61; Mismatches 172; Indels 107; Gaps 22;

```

```

: Sequence 2, Application US/07946497
: Patent No. 5506119
: GENERAL INFORMATION:
: APPLICANT: HERRLICH, Peter
: APPLICANT: PONTA, Helmut
: APPLICANT: GUENTHER, Ursula
: APPLICANT: MARZKU, Siegfried
: APPLICANT: WENZL, Achim
: TITLE OF INVENTION: VARIANT CD4 SURFACE PROTEINS, DNA
: TITLE OF INVENTION: SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
: TITLE OF INVENTION: AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 3000 K Street, N.W., Suite 500
: CITY: Washington, D.C.
: COUNTRY: USA
: ZIP: 20007-5109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/946,497
: FILING DATE: 19921109
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 16915/145
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)672-5300
: TELEFAX: (202)672-5399
: TELEX: 904136
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 503 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-07-946-497-2

Query Match 3.7%; Score 88.5; DB 1; Length 503;
Best Local Similarity 18.4%; Pred. No. 1.5;
Matches 57; Conservative 32; Mismatches 96; Indels 125; Gaps 14;

```


Db 213 ---DAFFGSLTATATPTWWSAHTKQNERQWNPINSPVLLQTT----- 258
QY 134 MLTKDAKOLEYAAKQTPPLSLSPDLDRNNPL---WSSRPNPMYVLPITMGKPNRS 209
Db 259 -----RMF-----DIDRNSTSAGENMTQEPQPF----- 283
QY 210 PNTPSHAKQPTNPFRAPELKFQYSVAKAEDLM-----GTDSDLMFGTQOS 259
Db 284 -NNHEYODEETPH-----ATSTWADPNSTTEBAATQKEWF-----EN 322
QY 260 HMOIFNGKN-----SRPRVHDYOPEIFLTOPVYSDLPW---DGKVRMIGWG 303
Db 323 EMQ---GKNPTPSDSHTVEGTASAHNNHPSQKMTTQSOEDVSGTDFDPDISHPMGG 379
QY 304 AVHSHNGESA 313
Db 380 HOTESKGHS 389
RESULT 11
US-08-961-083-2
; Sequence 2, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961.083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8512
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 666 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-083-2
Query Match 3.7%; Score 87.5; DB 4; Length 666;
Best Local Similarity 20.4%; Pred. No. 3;
Matches 87; Conservative 60; Mismatches 176; Indels 103; Gaps 24;

Db 70 ----SITLQOILKLTFFSTSDQITIS-----RKADENMLAIQLEKAKIKQELTVY 117
QY 142 TDPIFLMGNEKMLTK-----KDAKQLE-----YAAKQTPPLS-LSFDLDRNN 183
Db 118 INKVVYMSNGYMCQTAQNYKDKLNNLSLPOLALLAGMPQAPBNQDYPYSHPEAADRNN 177
QY 184 TPLMSRRPNPMYVLPITFMHGKPNRSPNPSHAKQFTNERRAP-----LKFOYS-VK 237
Db 178 LVL--SEMKNOGYI-----SAEQYERAVNPTITDGLQSLKSASNYPMYMDNYLKEVINQVE 231
QY 238 VAAEIDTDDSLDMFGYTO--QSH-WQIFNGKNSRPFVHDYOPEIFLTOPVYSDLPWD 294
Db 232 EETGYNLITGMGDYTYTNVQDEAKHLMIDLYNDEYAYPPDDELQ-----VASTYDVS-N 285
QY 295 GKVYRMIGMGAHVHNS-----GESAKLSRWBRAYLMAEMEKMLT-VMPRI-----W 340
Db 286 GKV-IAQLGARHQSSNVSGINQAVETNRDW-----GSTMKPTIDYAPALEYGYVDST 337
QY 341 GRIFKEGSGSDPDNDIILDY--YGY-GDYRFLYQLENKSNISGTAVYN-----PRSGKG 392
Db 338 ATIVHDEPYNYGTNTPVYNMDRGYFGNITLQYALQDSHNVPAVETLNKVGILNRAKTFLN 397
QY 393 ALQLDY 398
Db 398 GLGIDY 403
RESULT 12
US-08-373-134D-2
; Sequence 2, Application US/08373134D
; Patent No. 5780296
; GENERAL INFORMATION:
; APPLICANT: Kmiec, Eric
; APPLICANT: Holloman, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS TO PROMOTE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Penile & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/373,134D
; FILING DATE: January 17, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Friebel, Thomas E.
; REGISTRATION NUMBER: 29,258
; REFERENCE/DOCKET NUMBER: 7991-007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 781 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-373-134D-2
Query Match 3.6%; Score 86; DB 1; Length 781;
Best Local Similarity 20.7%; Pred. No. 5.5;


```

QY 107 HGEIPAVIKTRKSRIRLDEITIMQIKGKPOVYUOE--TDDPI---FLMGNEKMGMLTKKDAQ 162
Db 614 FTQAPSTVPKAT-----QITSTPEVLQSPSTPEVPFTRTLGAEBEITQTPBSA-- 662
QY 163 LEYAKOFTPLTSLFDDLRNNTPLMSSRP-----HN-----PMVLPL-----IFMHGK 205
Db 663 ---APEYVTRBSSSTIMPETAOSTPIPLASONPIRSSGCTGNTTERPRIYTPVOYTPPHQOKLTENK 719
QY 206 PNRSPNTPS--HE---AKOFTP--NEFRAPELPOVSVYKVA-----AEDIMGTD 248
Db 720 TLRSPYVSEEHKSTAESQRPDLDAIVKEKFSNDEVEYATATCVSYKSPRYVETIMNKVD 779
QY 249 -----SDLMFGYVQOOSHQJFNGKNS-----RPF----- 272
Db 780 LVYVMEDEISGNSPAGVFNSENK---OKOLYYRVTDGRTSVQJMLCISCTSHSPEBYCLFD 836
QY 273 -----RHVDYQPELFL---QPVYSDLPMDGKV-----RHIGMGAVHNHN 309
Db 837 TSLIAREKIDIPBELYFTSDPOTACITLPSGVVPREFEWSLNNVSLPEYLATATVVSHTA 896
QY 310 GE-----SAKLSRSM 319
Db 897 GOSTVWKKSSARAGEAW 912

```

RESULT 15
PCT-US96-03916-66

; GENERAL INFORMATION:

APPLICANT: Wild, Martha A.
APPLICANT: Cochran, Mark D.
APPLICANT: Cochran, Mark D.

TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
 TITLE OF INVENTION: AND USES THEREOF
 TITLE OF INVENTION:

NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS: ADDRESS

ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York

CITY: New York
STATE: New York
COUNTRY: U S A

COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE

```

COMPIER READABLE FORM:
MEDIUM TYPE: Floppy
COMPIER: IBM PC COM

```

COMPUER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS
SOFTWARE: RatonTrn Pol3200

```

SOFTWARE:  Patchlin release #1.23
CURRENT APPLICATION DATA:
APPLICATION NUMBER:  DCE/US06/03016

```

APPLICATION NUMBER: FC/

CLASSIFICATION:
PRIOR APPLICATION DATA:
ADDI TION NUMBER:

APPLICATION NUMBER: US 08/120,391
FILING DATE: 24-SEP-1993

```
;
; ALIEN/AGENT INFORMATION;
NAME: White, John P.
REGISTRATION NUMBER: 38 679
```

REGISTRATION NUMBER: 28,076
REFERENCE/DOCKET NUMBER: 39116-A
TELECOMMUNICATION INFORMATION:

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 301-0535

```

; TELEFAX: (212) 391-0323
; INFORMATION FOR SEQ ID NO: 66:
;     SEQUENCE CHARACTERISTICS:

```

SEQUENCE CHARACTERISTICS:

LENGTH: 985 amino acids

TYPE: amino acid

MOLECULE TYPE: protein

PCT-US96-03916-66

Query Match	3.68;	Score 86;	DB 5;	Length 985;
Best Local Similarity	19.48;	Pred. No. 8;		
Matches	73;	Conservative	47;	Mismatches 126;
				Indels 130;
				Gaps 20

```

OY 48 ELPIIDYOSATQSADPRANPLDEHEPELYT-TALENTKMLINCALNQDMLACYLTIV 106
Db 563 QTPSTVEKTTLLSSST-----EPALFTTOSAGTEATLTQTSASAPDTRFQSTETHF 61.33
OY 107 HGETPAVVKTKRSIRLDETTIMQTKGPROVYOE-TDPI---FLMONEKGMILTKKDAQ 162
Db 614 FTQAPSTVPKAT-----QFTSTEBEVLTOSPRSTPEVPFTKTLCAPEIOTPSA-- 66.22
OY 163 LEYAAKOFTPLSLSEDDLRNTPMLMSSRP-----HN-----PMYLP----IFMHK 205
Db 663 ---APEVYTRSSSTMPETAOSTPLASONPTSSGCTGHTNTERPRYVQTPHTOKLYENK 719
OY 206 PNRSPNTPS--HE--AKQFPR-NEFARPELKPQVSVKYKA-----AEDLIGTD 248
Db 720 TILSPFTVVSERHEMSTAESQTPRLLDVAIVEKFSNDGEVATATCVSVKSPRYRETNMKVD 77.9
OY 249 -----SDLMEGYQOOSHMOJFNKNS-----RPF----- 272
Db 780 LVDVWDELSGNSPAGVFNSSNKM---QKOLYKRTDRTVSQMLCISCTSHSPERYLFD 83.66
OY 273 ----RHVDYQPELFLT--QPVYSDLPWDCGV-----RMIGMCAVHNHN 30.9
Db 837 TSLIAREKDIDPELYFTSDPOTACTITLPSGVVPRFEWMSLNNYSLPEYLTATVVSHTA 89.66
OY 310 GE-----SAKLSRW 319
Db 897 GOSTVWKSSARAGEAW 912

```

Search completed: November 30, 2001, 14:17:51
Job time: 228 sec

2 . . 2

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:18:35 ; Search time 44.45 Seconds
(without alignments)
757.461 Million cell updates/sec

Title: US-09-787-083-8

Perfect score: 2363

Sequence: 1 MKVSLSTLTSLTILPCFAILA.....YNHEATSEFGVGLMNDMNGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	815	34.5	382	2	E81195 phospholipase A1, probable phospholip
2	815	34.5	409	2	H81831 outer membrane pho
3	373	15.8	286	2	B36971 phospholipase A1 (
4	366	15.5	329	2	D81279 outer membrane pho
5	358	15.2	289	2	A36971 phospholipase A1 (
6	357	15.1	289	1	P58CA1 phospholipase A1 (
7	357	15.1	289	2	E86069 outer membrane pho
8	321.5	13.6	289	2	C36971 phospholipase A1 P
9	246.5	10.4	355	2	C64582 phospholipase A1 P
10	240.5	10.2	355	2	H71930 probable phospholi
11	111.5	4.7	602	1	TYVPRR protein kinase (EC
12	108.5	4.6	800	2	A28003 cellulase (EC 3.2.
13	108.5	4.6	822	2	JT0611 cellulase (EC 3.2.
14	107.5	4.5	824	3	JC7532 cellulase (EC 3.2.
15	106.5	4.5	901	2	T20122 hypothetical prote
16	100.5	4.3	660	2	E83656 methionyl-tRNA syn
17	100.5	4.3	783	2	JC5467 cellulase (EC 3.2.
18	100.5	4.3	1658	2	T42642 phospholinositide 3
19	99.5	4.2	797	2	T46737 X-Pro dipeptidyl-P
20	99	4.2	5005	2	F82884 hypothetical prote
21	98.5	4.2	810	2	S49744 AMP deaminase (EC
22	98.5	4.2	838	2	A96557 AMP deaminase (EC
23	96	4.1	791	2	H96839 hypothetical recepto
24	96	4.1	4307	2	T20721 hypothetical prote
25	95.5	4.0	719	2	A42893 penicillin-binding
26	95	4.0	765	2	T35719 chitinase - Strept
27	95	4.0	888	2	S50801 AMP deaminase homo
28	94.5	4.0	397	2	A35136 cellulase (EC 3.2.
29	94.5	4.0	601	2	T26062 hypothetical prote

30	94.5	4.0	655	2	T26061 hypothetical prote
31	94.5	4.0	1310	2	T40135 probable involveme
32	94.5	4.0	1478	2	S20117 protein kinase BCK
33	93.5	4.0	564	2	T40777 ferlic reductase t
34	93.5	4.0	719	2	S28031 penicillin-binding
35	93.5	4.0	719	2	S28033 penicillin-binding
36	93.5	4.0	857	1	A41369 S-receptor kinase
37	93	3.9	1078	2	T18352 protein P120 - Myc
38	92.5	3.9	486	2	S30959 gene 14 protein -
39	92.5	3.9	719	2	S28034 penicillin-binding
40	92.5	3.9	719	2	S28032 penicillin-binding
41	92	3.9	324	2	B69521 hypothetical prote
42	92	3.9	470	2	T43675 cog-2 protein - Ca
43	91.5	3.9	796	2	JC7355 peroxisome prolif
44	91.5	3.9	1641	2	D82704 conserved hypothet
45	91	3.9	367	2	T24058 hypothetical prote

ALIGNMENTS

RESULT 1	phospholipase A1, probable NMB0464 [imported] - Neisseria meningitidis (strain MC58 s
E81195	
C:Species: Neisseria meningitidis	
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001	
C:Accession: E81195	
R:Retellin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, R.E.; Eisen, Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B. ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masiagnan, V.; Pizzoli, M. Science 287, 1809-1815, 2000	
A:Authors: Grandt, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R. A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58. A:Reference number: A81000; MUID:20175755	
A:Accession: E81195	
A:Status: preliminary	
A:Molecule type: DNA	
A:Residues: 1-382 <TEXT>	
A:Cross-references: GB:AE002403; GB:AE002098; NID:97225688; PIDN:AAF40901.1; PID:9722	
A:Experimental source: serogroup B, strain MC58	
C:Genetics:	
A:Gene: NMB0464	
Query Match	34.5%; Score 815; DB 2; Length 382;
Best Local Similarity	44.7%; Pred. No. 3.5e-58;
Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;	
OY 87	INCSALNODIMRLACVDTLVHGETPANI-----KTRSTRLDETIMGTI-KGRPOVYOE 140
DB 33	LOCALALTDNVTIRLACYDRIFAAQLPSSAGEGSEKAVLNLTEVRSLLDKGEVIVYEK 92
OY 141	TTPFIPLMGNEKMLTKKDAKOLEYAAKOPTPLSEFDLRNN-TPLMSRPHNPVLP 199
DB 93	GGAL-----PADSAGETRADITPTLSIMLDLKDNDRLGLGYREINPWLMP 139
OY 200	IFNHGKPNRSPNPSH-EAKQTPNEFRABELFOYSVKAAEDLMGTSDLMFGYTOQ 258
DB 140	LWVNNSPNVAFGSPTRTQVQEFKGOOKRAETKLOVSFKIADLEFKTRADLMFGYTOR 199
OY 259	SHMOIIN-GKNSRPFVHDYQPELITQPYISDLPMDGKVRMIGMCAVHNSNESAKLR 317
DB 200	SDMOIYNQGRKSAFRTWDKPEIFLTPQYKADLPFGGRILMLGAGVHNSQOSRPSR 259
OY 318	SNRRATLMGMEKNTLVMPRIWGRITFEKSGSQOPDNDPILDYGGVDFRYOLENKS 377
DB 260	SNRRATLMGMEKNTLVMPRIWGRITFEKSGSQOPDNDPILDYGGVDFRYOLENKS 377
OY 378	NISGTVRNPRSGKALQLDYVPLGKGISGYFOIFOGYGOGLIDYNHEATSEFGVGLMN 437
DB 318	NVYSVLRNPKTGYGAIEAAYTPRIKGLKGVVRFHGYGESLIDYNNHKGNGIGIGIMFN 377
OY 438	DNMGL 442

Query Match 10.2%; Score 240.5; DB 2; Length 355;
 Best Local Similarity 25.1%; Pred. No. 8, 9e-12;
 Matches 83; Conservative 44; Mismatches 109; Indels 95; Gaps 13;

QY 157 KDAKOLEYAKOFTPLSLFSDLRNNTPLMSSRPHNPMVYLPIFMHGKPNRSPNTSHE 216
 DB 69 KKYLNMDYLGTYFLPFYHSF-----TFIFQWHPHNP----- 102

QY 217 AKOTPEFAPRELKFOVSKVKAEDLWCTDSLMFGYTOOSHWOIFNGKNSRPFVHD 276
 DB 103 ---YORNEF-----KFOISFVYFVRHILWTKGLYLYTOTNPFQIYNDPOSAFMRMIN 154

QY 277 YOPFIFLTOPYPSDLPMDGKV---RMIGMAVHSNG-ESAKLSRSMNR----- 321
 DB 155 FMPFLIYVYPI-NKPPFGKIGNFSEIWMQHSNGVGAQCYQPFKEGKNPENOPFGQ 213

QY 322 -----AYLMAG-----MEMK-NLTVPRIMGRIFKEGSGSOPDD- 354
 DB 214 PVIKYDYNQKDVWGGCRSVSAGNALCFVLWMEKGGKIMVAYMPVY-----PYDQ 265

QY 355 -NPDIIDYGYGVRFY-----OLEKNSNIGTVRYRPNRSGKALQLDYVPLGKI 406
 DB 266 SNPDLIDMYGNAKIDYRGRHHEFLQYDIFQWYTD--RWHGAFRLGYTYRINPFV 323

QY 407 SGYFOIFOGYGQSLIDYNHEATSFVGMLN 437
 DB 324 GIYAQMFNGYGDGLYEYDFVSNRIGVGIRLN 354

RESULT 11

TVTRRR
 protein kinase (EC 2.7.1.37) raf - raf
 N:Alternate names: kinase-related transforming protein raf; raf proto-oncogene protein-s
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 11-Jun-1999
 C:Accession: B26126
 R: Ishikawa, F.; Takaku, F.; Nagao, M.; Sugimura, T.
 Mol. Cell. Biol. 7, 1226-1232, 1987
 A:Title: Rat c-raf oncogene activation by a rearrangement that produces a fused protein.
 A:Reference number: A26126; MUID:87172791
 A:Accession: B26126
 A:Molecule type: mRNA
 A:Residues: 1-602 <ISH>
 A:Cross-references: GB:M15428; NID:g206546; PIDN:AAA42002.1; PID:g206547
 C:Genetics:
 A:Gene: raf
 C:Superfamily: rat protein kinase raf; protein kinase homology
 C:Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; serine/threonin
 F:301-567/Domain: protein kinase homology <KIN>
 F:309-317/Region: protein kinase ATP-binding motif
 F:329/Active site: lys #status predicted
 F:453/Binding site: phosphate (Ser) (covalent) (by autophosphorylation) #status predicted

Query Match 4.7%; Score 111.5; DB 1; Length 602;
 Best Local Similarity 19.3%; Pred. No. 0.53;
 Matches 93; Conservative 71; Mismatches 174; Indels 145; Gaps 23;

QY 23 QAAQVNPVAFVDEVRSENDLGDNELPIDVQSASTPTANPLDEHPELTTALEN 82
 DB 81 ELEKLNNOVAVLTKKNELETAODNRIGI-----QSOFTRAKELEKEKDLRT---N 131

QY 83 KTMILNCSALNODIMRLACYDTLVHGETPAVYIKTKRSIRIDETIWO--TIKGPQVYOE 140
 DB 132 ERLSOEVEYLETEDVKRL--NEKTKESNT---TKGELQKIDELQASDVTVYKREKLEOE 186

QY 141 -----TTDPFILMGNEK-----MLTKDAOLEYAAQFTPLS 174
 DB 187 KELLHONNSMLNELLKTKTTELLALGREGKEIILELCTKENKEEDAIRSHSESASPSA 246

QY 175 LSFDLDRNNTPLMSSRPHNPMVYVPIFMHGKPNRSPNTSHEAKOFTPNERRAELKFOV 234
 DB 247 LS-SSPNNLSPTGWSQKPTP-----VPAQREAPGSGTQEKKNKIRRGGRDSYYWEI 298

QY 235 SVKYAAADLNGTSDLMFGYTOOSH-----QIFNCKNSRPFVHDYQPEI----- 281
 DB 299 EASEVMSTRIGSGS--FGTVYRGKMGDVAVKLLKYVDLPPEQLQAFRNEVAVALRTR 355

QY 282 -----FLTO-----PYSDL-PWDGKYRMI-----GKGAVAH 307
 DB 356 HVNILLFMGYMTKDNLAIVTQWCSSSLXKHLHVOETFFQMFQDLIDIAKOTAGNDYLHA 415

QY 308 SNGESAKLSRSMNRAIYLAGMEK---NLTVPRIMGRIFKEGSGS----- 351
 DB 416 KNIHTRDKS--NNIFLEHGLTVKIGDEGLATVSRW-----SGSQOVEDQPTGSVLM 466

QY 352 -----PDDNPDL--DYGYGVRFYLOLEKNSNIGTVRYRPNRSGKALQLDYVPL 402
 DB 467 APEVIRMODNPFSPQSDVSYGIV--LYEL-----MTGELPYSHNNRQI---IEMV 515

QY 403 GKG 405
 DB 516 GRG 518

RESULT 12

A29003
 cellulase (EC 3.2.1.4), alkaline - Bacillus sp.
 N:Alternate names: endo-1,4-beta-glucanase
 C:Species: Bacillus sp.
 C:Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 21-Jul-2000
 C:Accession: A29003
 R: Fukumori, F.; Kudo, T.; Narahashi, Y.; Horikoshi, K.
 J. Gen. Microbiol. 132, 2329-2335, 1986
 A:Title: Molecular cloning and nucleotide sequence of the alkaline cellulase gene fro
 A:Reference number: A29003; MUID:87085443
 A:Accession: A29003
 A:Molecule type: DNA
 A:Residues: 1-800 <FKR>
 A:Cross-references: GB:D00066; GB:N00066; NID:g216223; PIDN:BA00045.1; PID:g216224
 A:Experimental source: strain 1139
 C:Function:
 A:Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as
 A:Pathway: cellulose degradation
 C:Superfamily: Bacillus sp. alkaline cellulase; Thermotoga xylanase A amino-terminal
 C:Keywords: glycosidase; hydrolyase; polysaccharide degradation
 F:565-726/Domain: Thermotoga xylanase A amino-terminal repeat homology <TXA>

Query Match 4.6%; Score 108.5; DB 2; Length 800;
 Best Local Similarity 19.9%; Pred. No. 1.4;
 Matches 96; Conservative 54; Mismatches 159; Indels 173; Gaps 25;

QY 8 LTLILPCFALIAIQAAQVNPVAFVDEVRSEND--LGODNELPIDVQSASTPT 64
 DB 10 LISSILVILLST-----PPTALAEGRNTRDNKRLHNDVYKRPSEAGALQLOEYDG 64

QY 65 ANPL-DEHEPLY-----TTLLENKTMILNCSAL-----NODIMRLACY-DTLVHGET 110
 DB 65 QMTLVDOGEKTIQLRGMSTHGLQWPELINDNAYKALANDWESNMIRLAMYGVGENGYSN 124

QY 111 PAVIKTKRSIRIDETI-----WQT-----IKGKPOV 137
 DB 125 PELIKSRVYIKGIDLAITEMDMVYIYDWHYHAGCDRDPYVACAEDFFRIALALYRPNPHIT 184

QY 138 YQETTD-----IFLMGNEKMLTKKDAKOLEYAAKQFTPLSLFSDLRN---NTP 186
 DB 185 YELANEPSSNNNGAGIYNNNEGMNAVYK-----EYADPIVEMLRDSGNADNIIIVGSPN 239

QY 187 WSSRP-----HNPVYLPPIF--MHGKPNNS--PNTSHEAKOFTPNERRAELK 232
 DB 240 WQSRDLAANDPIDDHHTMYTVHFYTSASHASTSYRPTNBSRGVMSYTRVA---L 295

QY 233 QVSQYKVAEDLWCT-----SDSLMFGYTOO-----SHWOIFNGKNS-----R 270
 DB 296 ENGVAVPAITE--WGTSGANGCGGPFDEADVWIEFLNENISMANNSLTN-KNEVSGAFT 352

This Page Blank (uspto)


```

OY 226 RAPELKFQVSVKKAEDLWGTDSLMFGYTOOSHQIFNGKNSRPFVHYDYOPELFLQ 285
    ++++++
DB 80 RKEVEFQSLAPFLMRGILGDSNLSGASYQKSMWOLSKSKSSPFRFETNYPQLFLGF 139
OY 286 PV-YSLDPMWGVKRMIGMCAVHHSNGESAKLSRNNRAYLMAEMKNLTVPRINGRI 344
    ++++++
DB 140 ATDYRFAGW--TLRDIEMGYNHDSNGRSDPTSRNRLYLRLMAONGNMLVEVKPM--Y 194
OY 345 KEGSGSQPDNDPILIDYGGVDFLYOLENKSINIGTYVYRNSRGKALQLDYVPLGK 404
    ++++++
DB 195 VVGS---TDDNDPITKYMGYRLKAVGYQL-GEAILSAQGYNNMTNGGAEGLSVSIPITK 250
OY 405 GISGTFQIFQGGQSLIDYVNEATSFVGLMLND 438
DB 251 HVRAATQIYSGYGESLIDYVNFQTRVGVGLMLND 284

RESULT 2
PAL-SALTY
ID PAL-SALTY STANDARD; PRT; 289 AA.
AC P37442;
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT
DE PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-
DE ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).
GN PLDA.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94131966; PubMed=8300539;
RA Brok R.G.P.M., Brinkman E., van Boxtel R., Bekkers A.C.A.P.,
RA Verheij H.M., Tommassen J.;
RT "Molecular characterization of enterobacterial plda genes encoding
RT outer membrane phospholipase A."
RT J. Bacteriol. 176:861-870(1994).
CC -1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE
CC A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
CC LOCATED THERE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X76900; CA54222.1; -
DR PIR: A36971;
DR PIR: S40131; S40131.
DR StyGene; SG10302; PLDA.
DR InterPro; IPR003187; PLA1.
DR Pfam; PF02253; PLA1; 1.
KW Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium.
FT SIGNAL 1 20
FT CHAIN 21 289 BY SIMILARITY.
FT ACT_SITE 164 164 PHOSPHOLIPASE A1.
FT ACT_SITE 164 164 BY SIMILARITY.
SQ SEQUENCE 289 AA; 33003 MW; D5ACFE533F704DBA CRC64;

```

```

Best Local Similarity 39.3%; Pred. No. 1,2e-21;
Matches 84; Conservative 30; Mismatches 90; Indels 10; Gaps 4;

OY 226 RAPELKFQVSVKKAEDLWGTDSLMFGYTOOSHQIFNGKNSRPFVHYDYOPELFLQ 285
    ++++++
DB 83 RKEVEFQSLAPFLMRGILGDSNLSGASYQKSMWOLSKSKSSPFRFETNYPQLFLGF 142
OY 286 PV-YSLDPMWGVKRMIGMCAVHHSNGESAKLSRNNRAYLMAEMKNLTVPRINGRI 344
    ++++++
DB 143 ATDYRFAGW--TLRHVEMGYNHDSNGRSDPTSRNRLYLRLMAENGNNMLVEVKPMYVI- 199
OY 345 KEGSGSQPDNDPILIDYGGVDFLYOLENKSINIGTYVYRNSRGKALQLDYVPLGK 404
    ++++++
DB 200 -----GSTDNDPITKYMGYRLKIGYHL-GEAVLSAKGYNNMTNGGAEGLSVSIPVTK 253
OY 405 GISGTFQIFQGGQSLIDYVNEATSFVGLMLND 438
DB 254 HVRLTYQVYSGYGESLIDYVNFQTRVGVGLMLND 287

RESULT 3
PAL-ECOLI
ID PAL-ECOLI STANDARD; PRT; 289 AA.
AC P00631;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT
DE PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-
DE ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).
GN PLDA OR B3621 OR Z5342 OR ECS4751.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85157492; PubMed=6397464;
RA Homma H., Kobayashi T., Chiba N., Karasawa K., Mizushima H., Kudo I.,
RA Inoue K., Ikeda H., Sekiguchi M., Nojima S.;
RT "The DNA sequence encoding plda gene, the structural gene for
RT detergent-resistant phospholipase A of E. coli."
RT J. Biochem. 96:1655-1664(1984).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RC MEDLINE=92358234; PubMed=1379743;
RA Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;
RT "Analysis of the Escherichia coli genome: DNA sequence of the region
RT from 84.5 to 86.5 minutes."
RT Science 257:771-778(1992).
RN [3]
RP REVISION TO 14-15.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RT Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=O157:H7 / EDL933 / ATCC 700927;
RC MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."

```

Query Match

15.2%; Score 356; DB 1; Length 289;

DR	PIR: A22133; PSECA1.
DR	PIR: S30711; S30711.
DR	Ecogene: EG10738; plda.
DR	InterPro: IPR003187; PLA1.
DR	Pfam: PF02253; PLA1, 1.
KW	Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium;
KW	Complete proteome.
FT	SIGNAL 1 20
FT	CHAIN 21 289
FT	ACT_SITE 164 164
FT	MUTAGEN 172 172
FT	CONFLICT 14 15
FT	CONFLICT 30 33
SO	SEQUENCE 289 AA; 33163 MW; A688AD32AA60F218 CMC64;
Query Match	15.1%; Score 357; DB 1; Length 289;
Best Local Similarity	39.3%; Pred. No. 1.5e-21;
Matches	84; Conservative 29; Mismatches 91; Indels 10; Gaps 4;
QY	226 RAPELFQVSVKVKAAEDLMTDSLMEGYTQOSHQIFNGKNSRPFVHDYQPEIFLTQ 285
DB	83 RKDEYKFDLSIAFLPMRGILGPNISYLGASTQKSWMQLNSSESPFRETNEPOLFLGF 142
QY	286 PV-YSDLPWDGKVRVIGAGVAHSHSGESAKLSRSNNRATYLAAGMKWKLTYMPRIWRGRI 344
DB	143 ATDYEFAGM-TLRVEMGYNDHNSGRSDPTSRSWNRLYTRLAENGWMLVEVKPMYYV- 199
QY	345 KEGSSQSDDDNDIDYGYGVDVRFYLYOLENKSNTSGTVRYNPRSGKALOLDYVYPLGK 404
DB	200 -----GNTDDNPDTIKYMGITYOLKGYHL-GDAVLSAAGOTNMWNTGYGAELGLSTPYTK 253
QY	405 GISGYSQIFQGYGQSILIDYHNHATSGVGYGLMND 438
DB	254 HVRLTQYVSGSGESLIDYNFMTQVGVGMND 287
RESULT 4	
PAL_PROVU	STANDARD; PRT; 289 AA.
ID	PAL_PROVU
AC	P37447;
DT	01-OCT-1994 (Rel. 30, Created)
DT	01-OCT-1994 (Rel. 30, Last sequence update)
DT	01-NOV-1997 (Rel. 35, Last annotation update)
DE	PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETURGENT-RESISTANT
DE	PHOSPHOLIPASE A1 (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-
DE	ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).
GN	PLDA.
OS	Proteus vulgaris.
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC	Proteus.
OX	NCBI_TaxID=585;
RN	[1]
RP	SEQUENCE FROM N. A.
RX	MEDLINE=94131966; PubMed=8300539;
RA	Brok R.G.P.M., Brinkman E., van Bostel R., Bekkers A.C.A.P.,
RA	Verheij H.M., Tommassen J.;
RT	"Molecular characterization of enterobacterial plda genes encoding
RT	outer membrane phospholipase A.,"
RL	J. Bacteriol. 176:861-870(1994).
CC	-1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE
CC	A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
CC	-1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.
CC	-1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.
CC	-1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
CC	-1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
CC	LOCATED THERE.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its


```

RESULT 8
GUN_PAPEO STANDARD: PRT: 397 AA.
AC P23548:
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE ENDOGLUCANASE (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE) (CELLULOSE).
OS Paenibacillus polymyxa (Bacillus polymyxa).
OC Bacteria: Firmicutes: Bacillus/Clostridium group:
CC Bacillus/Staphylococcus group: Paenibacillus.
OX NCBI_Taxid=1406;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90170877; PubMed=2307659;
RA Baird S.D., Johnson D.A., Seligy V.L.;
RT "Molecular cloning, expression, and characterization of
  endo-beta-1,4-glucanase genes from Bacillus polymyxa and Bacillus
  citriculus."
RL J. Bacteriol. 172:1576-1586(1990).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
  LINKAGES IN CELLULOSE.
CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY A (FAMILY 5 OF GLYCOSYL
  HYDROLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
CC -----
DR EMBL: M33791; AAA22631.1; -
DR PIR: A35136; A35136.
DR HSSP: P54383; 1ECB.
DR InterPro: IPR001547; Glyco_hydro_F5.
DR Pfam: PF00150; cellulase: 1.
DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; 1.
KW Cellulose degradation; Hydrolase; Glycosidase.
FT ACT_SITE 194 194 PROTON DONOR (BY SIMILARITY).
FT ACT_SITE 317 317 NUCLEOPHILE (BY SIMILARITY).
SQ SEQUENCE 397 AA; 44357 MW; B9C2E802C04F0A2A CRC64;

```

```

Query Match 4.0%; Score 94.5; DB 1; Length 397;
Best Local Similarity 25.7%; Pred. No. 3;
Matches 46; Conservative 21; Mismatches 61; Indels 51; Gaps 10;

```

```

QY 301 GNGAVHNGESAKLSRSNRAVILMAGMEKMLTVPRIMGR-----IFKEG----- 347
DB 43 GKRIYDESKKEAFNGLNM-----FGLETPNVT-LHGILSRKMDMDLQVKKEGYNLR 95
QY 348 -----SGSQPDD-----NPDILDYGYGDVRFVYLENKSNTSGTV---RYNPRS 389
DB 96 LPYSNOLFSSRPSIDYKKNPDV---GLNPIQIMDLIKERAGRGIIILDRHRPS 152
QY 390 GNGAQLDLV--YPLCKGISGYEQIFQGY-----GSLIDYNHNAITFSGVGLMLNDW 439
DB 153 G-GQSELWYTSYPSRWISDMKMLADRYKKNNTYIGADLHNEPHQASMGCGNASTDW 210

```

```

RESULT 9
BCK1_YEAST STANDARD: PRT: 1478 AA.
AC Q01389; P32894;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SERINE/THREONINE PROTEIN KINASE BCK1/SLK1/SSP31 (EC 2.7.-.-).
GN BCK1 OR SLK1 OR SSP31 OR LMS3 OR YJL095W OR J0906.

```

```

OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_Taxid=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92186847; PubMed=1545797;
RA Costigan C., Gehring S., Snyder M.;
RT "A synthetic lethal screen identifies SLK1, a novel protein kinase
  homolog implicated in yeast cell morphogenesis and cell growth."
RL Mol. Cell. Biol. 12:1162-1178(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92104496; PubMed=1840547;
RA Irie K., Araki H., Oshima Y.;
RT "A new protein kinase, SSP31, modulating the SMP3 gene-product
  involved in plasmid maintenance in Saccharomyces cerevisiae."
RL Gene 108:139-144(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=EG123;
RX MEDLINE=92107166; PubMed=1729597;
RA Lee K.S., Levin D.E.;
RT "Dominant mutations in a gene encoding a putative protein kinase
  (BCK1) bypass the requirement for a Saccharomyces cerevisiae protein
  kinase C homolog."
RL Mol. Cell. Biol. 12:172-182(1992).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C;
RX MEDLINE=95176706; PubMed=7871887;
RA Miosga T., Boles E., Schaaff-Gerstenschlaeger I., Schmitt S.,
  Zimmermann F.K.;
RT "Sequence and function analysis of a 9.74 kb fragment of
  Saccharomyces cerevisiae chromosome X including the BCK1 gene."
RL Yeast 10:1481-1488(1994).
RN [5]
RP SEQUENCE OF 602-1104 FROM N.A.
RA Cusick M.E.;
RT Submitted (XXX-1992) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: SERINE/THREONINE PROTEIN KINASE INVOLVED IN A SIGNAL
  TRANSDUCTION PATHWAY THAT PLAY A ROLE IN YEAST CELL MORPHOGENESIS
  AND CELL GROWTH. THIS PATHWAY SEEMS TO START BY SMP3; THEN
  INVOLVE THE KINASE PC1 THAT MAY ACT ON THIS KINASE. BCK1 PROBABLY
  PHOSPHORYLATES MKR1 AND MKR2 WHICH THEMSELVES PHOSPHORYLATE THE
  MPK1 KINASE.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
CC -----
DR EMBL: M84389; -; NOT_ANNOTATED_CDS.
DR EMBL: D10389; BAA01226.1; -
DR EMBL: X60227; CAA42788.1; -
DR EMBL: X77923; CAA54896.1; -
DR EMBL: Z49370; CAA89389.1; -
DR EMBL: Z49369; CAA89388.1; -
DR EMBL: M88604; AAA21179.1; -
DR PIR: S20117; S20117.
DR PIR: S22285; S22285.
DR PIR: J01118; J01118.
DR PIR: J01432; J01432.
DR HSSP: P24941; 1A01.
DR SCD: S0003631; BCK1.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_kin_actsite.

```

DR pfam: PF00069; kinase; 1.
 DR SMART: SM00220; S_TKC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW Transferrase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
 KM ATP-binding; Phosphorylation.
 FT DOMAIN 1175 1440
 FT NP_BIND 1181 1189 ATP (BY SIMILARITY).
 FT BINDING 1204 1204 ATP (BY SIMILARITY).
 FT ACT_SITE 1303 1303 BY SIMILARITY.
 FT MOD_RES 1134 1134 PHOSPHORYLATION (BY PKC) (POTENTIAL).
 FT VARIANT 1119 1119 T -> P (IN BCK1-19; ACTIVATION).
 FT VARIANT 1120 1120 I -> K (IN BCK1-11; ACTIVATION).
 FT VARIANT 1120 1120 I -> T (IN BCK1-16; ACTIVATION).
 FT VARIANT 1146 1146 G -> V (IN BCK1-10; ACTIVATION).
 FT VARIANT 1174 1174 A -> P (IN BCK1-20; ACTIVATION).
 FT CONFLICT 59 59 F -> I (IN REF. 3).
 FT CONFLICT 79 79 E -> V (IN REF. 2).
 FT CONFLICT 264 264 A -> P (IN REF. 3).
 FT CONFLICT 279 279 N -> I (IN REF. 3).
 FT CONFLICT 703 714 RYPPSYDYDR -> STPKPRVITMTE (IN REF. 3).
 FT CONFLICT 795 795 S -> A (IN REF. 3).
 FT CONFLICT 802 802 L -> V (IN REF. 3).
 FT CONFLICT 808 808 A -> S (IN REF. 3).
 FT CONFLICT 903 903 T -> N (IN REF. 3).
 FT CONFLICT 919 919 T -> N (IN REF. 3).
 FT CONFLICT 960 962 ADA -> RDR (IN REF. 5).
 FT CONFLICT 1086 1104 RPVPDSSEYEFIDGLNCK -> VPIAHNTSYRMDLVKIN
 H (IN REF. 5).
 SQ SEQUENCE 1478 AA; 164194 MW; D586C3A497A5BB33 CRC64;

Query Match 4.0%; Score 94.5; DB 1; Length 1478;
 Best Local Similarity 22.5%; Pred. No. 19;
 Matches 98; Conservative 51; Mismatches 156; Indels 131; Gaps 24;

OY 36 EYRSENDLGQNDLPL--DVQSAFQASTDTANPLDEHEPELYTALENK--TMLINCS 90
 DB 950 DIFEENDLTFADAPPFDDSDDDSSDDIWSKKKTAPE--TNNENKKDEKSDNS 1006
 OY 91 ALNODIMRLACYDTLVHGEFTPAVITKRSIRLDETIMOTIKGPOVVOYETDPIRLMN 150
 DB 1007 THSDELF--YDS--QTQKMERK-----MTFRSPPEVVO-NLEKFFRRAN 1047
 OY 151 EKGMLTRKDAQLEVAQKQFPLSLFDDLRNNTPLMSSRPNNPVVLPIFMHGKPNRS- 209
 DB 1048 LDKPIT-----EGIASPTSPKSLDSLSPKNVASSRTEPSPRPV-----PDSY 1094
 OY 210 -----PNTPSHEAKQFTNERRA-----PELAKQVSYKKAAL-DLMGTDSDIM 252
 DB 1095 EFIQDGLNGKNKPLNOAK--TEKRRTKRTITTAHEASLARKNSVKLKRONTKMWGT---RM 1149
 OY 253 FGYTQASHMOJENGNKNSRFVHDYQPELFLQPYSDLPMDGKVMIMGA-----V 305
 DB 1150 VEVTEHNHWSIKAKNSK-----GEYKEPAM-MKGEMIKSGSGAVYLCL 1193
 OY 306 HHSNSESAAKLRS-----SMNRATLMAGMEKKNLTVMPRIRGIRTEKSGSGQPDNDPI 358
 DB 1194 NVTTEGMAVVKQVEPKYSSQNEALLS-----TV-----EALNSEVSTLKLDHLMNI 1240
 OY 359 LDYVGYGVARFLYOLENSNIGSYVRYNPRSGKALQLDYVPLGKIGSYQIPIQGYGQ 418
 DB 1241 VOYIGF-----ENKNNTY-----SLFLEYV--AGSGVSLIRMYGRFDE 1277
 OY 419 SLIDYNHEATSPGVGL 434
 DB 1278 PLI--KHLFTVYLKGL 1291
 RESULT 10
 SRK6_BRAOL STANDARD: PRT: 849 AA.

AC 009092;
 DT 01-FEB-1995 (rel. 31, Created)
 DT 01-FEB-1995 (rel. 31, Last sequence update)
 DT 01-NOV-1995 (rel. 32, Last annotation update)
 DE PUTATIVE SERINE/THREONINE KINASE RECEPTOR PRECURSOR (EC 2.7.1.37)
 DE (S-RECEPTOR KINASE) (SRK).
 GN SRK6.
 OS Brassica oleracea (Cauliflower).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Brassica.
 ON NCBI_Taxid=3712;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. S656; TISSUE=Stigma;
 RC MEDLINE=92020942; PubMed=1681543;
 RX Stein J.C., Howlett B., Boyes D.C., Nasrallah M.E.;
 RT "Molecular cloning of a putative receptor protein kinase gene encoded
 at the self-incompatibility locus of Brassica oleracea";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:8816-8820(1991).
 CC -1- FUNCTION: INVOLVED IN SPOROPHYTIC SELF-INCOMPATIBILITY SYSTEM
 CC (THE INABILITY OF FLOWERING PLANTS TO ACHIEVE SELF-
 CC FERTILIZATION), PROBABLY ACTING IN COMBINATION WITH S-LOCUS-
 CC SPECIFIC GLYCOPROTEINS. INTERACTION WITH A LIGAND IN THE
 CC EXTRACELLULAR DOMAIN TRIGGERS THE PROTEIN KINASE ACTIVITY OF THE
 CC CYTOPLASMIC DOMAIN.
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN - ADP + A PHOSPHOPROTEIN.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY IN THE PISTIL AND ANTHR.
 CC -1- POLYMORPHISM: THERE ARE A NUMBER OF DIFFERENT S ALLELES IN
 CC B.OLERACEA, POSSIBLY PROVIDING THE RECOGNITION SPECIFICITY.
 CC -1- SIMILARITY: THE EXTRACELLULAR DOMAIN IS SIMILAR TO S-LOCUS
 CC GLYCOPROTEINS OF BRASSICA, WHILE THE INTRACELLULAR DOMAIN IS
 CC A SER/THR-PROTEIN KINASE RELATED TO RAF KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: M76647; AAA33000.1; ALT-TERM.
 DR HSSP; P11362; IFCI.
 DR InterPro: IPR001480; B_lectin.
 DR InterPro: IPR000719; Euk-kinase.
 DR InterPro: IPR003609; Pan-app.
 DR InterPro: IPR002280; Ser_thr_kin_actsite.
 DR InterPro: IPR000858; Slocus_glycop.
 DR pfam: PF00069; kinase; 1.
 DR pfam: PF00954; S_locus_glycop; 1.
 DR SMART; SM00108; B_lectin; 1.
 DR SMART; SM00473; PAN_AP; 1.
 DR SMART; SM00221; STYKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW Transferrase; Serine/threonine-protein kinase; Signal; ATP-binding;
 KW Transmembrane; Receptor; Glycoprotein; Self-incompatibility.
 FT SIGNAL 1 32
 FT CHAIN 33 849
 FT DOMAIN 33 446
 FT TRANSMEM 447 466
 FT DOMAIN 467 849
 FT DOMAIN 528 779
 FT NP_BIND 534 542
 FT BINDING 556 556
 FT ACT_SITE 653 653
 FT CARBOHYD 47 47
 FT CARBOHYD 120 120
 FT CARBOHYD 196 196
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 389 389 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 849 AA: 97231 MW: 7E156059EEDDF4370 CRC64;

Query Match 4.0%; Score 93.5; DB 1; Length 849;
 Best Local Similarity 21.9%; Pred. No. 10;
 Matches 78; Conservative 50; Mismatches 107; Indels 121; Gaps 24;

QY 105 LNHGEPFAVTKRSLDEITWOTIKGPQVYVQETDPIELMG----NKK---GMLTK 157
 Db 24 LHPALSIYNT-----LSSTESLTSSNKTLY--SPGSEFVEGFRTSRNWTGLMWYK 75
 QY 158 K-DAKOLEYAKOFTPLSLFSD-----LDRNNTPLMSSRPNNPMVLEPIFMHK 205
 Db 76 KVSRRYVWVANRNPJSMNIGTLKISGNNLVLDHKNKPYMTN-----LIRGN 125
 QY 206 PNRSP-----NTPSHEAKOFT-----PNEFRAPELKFOVYKVAEDLW 245
 Db 126 -ERSPVVAELIANGNFVWRDSSNNDASEYLMQSEFDYPTDTLLPEMKLGYNLKT----- 177
 QY 246 GTDSDLMEFGYTQOSHMOIFNCKNSRPRVHDYQ-----PEIFLQ---PYVSDLPMDGK 296
 Db 178 GLNFFL-----TSMSSDDPSSGNF---SYKLETSLEPEYLSRENEFPHRSGPMWG- 226
 QY 297 VRMIGMAVH-----HSNGESAKLSRWNRAY-----LMAGMEKMLTYMP--RI 339
 Db 227 IREFGIEDQKLSYMYNFIENNEVAYTFMTNNSPRTLSSEGFQGLTWYPSIRI 286
 QY 340 WGRIFKSGSGQPDNDPILDY-----YGYDVRFLYOLEKNSNIGTVR-YNPRS 389
 Db 287 WNRFW-----SSPYD-POCDTYIMCGPYAYCDV-----NTSPVCNIOGFNPRN 329

RESULT 11
 VG14_BPML5 STANDARD; PRT; 486 AA.
 AC 0052220;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE GENE 14 PROTEIN (GP14).
 GN 14.
 OS Mycobacteriophage L5.
 OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae.
 OX NCBI_TaxID=31757;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93211282; Pubmed=8459766;
 RA Hatfull G.F., Sarkis G.J.;
 RT "DNA sequence, structure and gene expression of mycobacteriophage L5:
 RT a phage system for mycobacterial genetics.";
 RL Mol. Microbiol. 7:395-405(1993).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: Z18946; CAA79390.1; -
 DR PTR: S30959; S30959; -
 SQ SEQUENCE 486 AA: 53757 MW: 4BF3734A6BB11BF5 CRC64;

Query Match 3.9%; Score 92.5; DB 1; Length 486;
 Best Local Similarity 21.0%; Pred. No. 5.8;
 Matches 99; Conservative 69; Mismatches 173; Indels 131; Gaps 24;

QY 21 IQQAQA-VPNVAVDEVRSN-----DLGDNELPIDVQATQSASTDTANPDEHEPE 74
 Db 56 MOQLAHGVPRLYVDAVRAQAVEGRFLGDADADELMQWQ-----ANNIDIAEP 109
 QY 75 LYTALEKMTKLINCSA-----LNODIMLACYDLVHGE-TPAVIKTRSIIRD 123
 Db 110 GYDAVYHGRSEFTISRPDPDLGMDQNVPIIIVE-PPTMHAIEDPRIKRSKAI-- 166
 QY 124 ETIWQIKGKPOVYVQETDPIELMGKMLTKKDAKOLEYAKOFTPL-SLSFDLDRN 182
 Db 167 -----VAYDKR-----GNE-----IQAAATLYPMETIGMFRADG 195
 QY 183 NTPLMSSRPNNPMVLEPIFMHKRNRSNTPSHEAKOFTNEFRAPELKFOVYKVAEAE 242
 Db 196 EWAEMFNYPHG-LGVVPVY--PLPNRRLSDYQTSSETPELRSMTQAAAILMLQATA 252
 QY 243 DLMEGTDSDLME-----GYTQOSHMOIFNCKNSRPRVHDYQPEI-----F 282
 Db 253 ELKGVPOQLFICIRPEELGVSETGOTFLDYLRIALFEDEAGKIQOFSAAELANFTNA 312
 QY 283 LTQ-----PYVSDLPMDGKVRMIGMAVHSHNGES-----AKLSRWNRAYLMAGMEKN 332
 Db 313 LDQIAKQVAAVTGLP-----PQYLSTAADNPASAEAIRAESRLIKYERKNLMGGAEE 368
 QY 333 LTVMPRIMGRIFKSGSGSQPDNDPILDYGYGCVRFYOLEKNSNIGTVRNPBSGK 392
 Db 369 AM-----RIAYRIMKG-----DVPPDM-----RME-----TYWRDPSTPT 401
 QY 393 ALQD-----YVYPLGKI--SGYFOIFQGYG-----QSLIDYHNEATSFVGILM 435
 Db 402 AAKADAATKLINGOGVTPRRARIDMGYSYKEREMRMBEAEAMOLGILL 453

RESULT 12
 BAC2_MOUSE STANDARD; PRT; 716 AA.
 ID BAC2_MOUSE
 AC P97303;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE TRANSCRIPTION REGULATOR PROTEIN BACH2 (BTB AND CNC HOMOLOG 2).
 GN BACH2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BAIB/C;
 RX MEDLINE=97042438; Pubmed=8887638;
 RA Oyake T., Itoh K., Motohashi H., Hayashi N., Hoshino H., Nishizawa M.,
 RA Yamamoto M., Igarashi K.;
 RT "Bach proteins belong to a novel family of BTB-basic leucine zipper
 RT transcription factors that interact with Mafk and regulate
 RT transcription through the NF-E2 site.";
 RL Mol. Cell. Biol. 16:6083-6093(1996).
 CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS AS REPRESSOR OR
 CC ACTIVATOR. BINDS, IN-VITRO, TO NF-E2 BINDING SITES. PLAY IMPORTANT
 CC ROLES IN COORDINATING TRANSCRIPTION ACTIVATION AND REPRESSION BY
 CC MAFK.
 CC -1- SUBUNIT: HETERODIMER OF BACH2 AND MAFK.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSION RESTRICTED TO MONOCYTES AND
 CC NEURONAL CELLS.
 CC -1- SIMILARITY: BELONGS TO THE BTB/POZ DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 BTB/POZ DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

```

CC      or send an email to license@lsb.slb.ch).
CC      EMBL; D86604; BAA13138.1; -.
DR      HSSP; P05412; IPOS.
DR      MGD; MGI:894679; Bach2.
DR      InterPro; IPR000210; BTB_POZ.
DR      InterPro; IPR001871; bZIP.
DR      Pfam; PF00651; BTB; 1.
DR      Pfam; PF00170; bZIP; 1.
DR      SMART; SM00338; BRLZ; 1.
DR      SMART; SM00225; BTB; 1.
DR      PROSITE; PS50097; BTB; 1.
DR      PROSITE; PS00036; bZIP_BASIC; 1.
KW      Transcription regulation; Activator; Repressor; DNA-binding;
KW      Nuclear protein.
FT      DOMAIN 37 103 BTB.
FT      DOMAIN 162 168 POLY-GLU.
FT      DNA_BIND 527 542 BASIC MOTIF.
FT      DOMAIN 550 572 LEUCINE-ZIPPER.
SQ      SEQUENCE 716 AA; 78935 MW; 9132B3731AE24333 CRC64;

Query Match
Best local Similarity 22.8%; Score 92.5; DB 1; Length 716;
Matches 93; Conservative 39; Mismatches 127; Indels 149; Gaps 21;

QY      38 RSNNDLG-----QDNELPIDVOSATQASTD-----TANPLDEH-----EPE 74
DB      151 RPOEDGNSAGEEEETWDSFETARMACATDMLDPDISFEATAIPIVAKEBALPES 210
QY      75 LVYTALENKTMILNCALNO---DIMRLAC---YDTLVHG----- 108
DB      211 VPTDTKENSEK---GALTDPYPRKYKQQLACTKNVYSAPSHGSGFASFSSEDSPOGSLK 266
QY      109 -----ETPAVITKTKSRIL---DETIWQTIKGPQVVOYETTDPI----- 145
DB      267 PGLPMGQIKSEPESEETESITLCLSGDET---DIKDRGDVDMRKQSPARTPTSTR 323
QY      146 -----FLMGNEKGMILTKKAKOLEYAAKOPTPLSLF----- 177
DB      324 GAACLDRSRSVSSPCLRLSIFGITKGV---ESTGLPSTSQ---PLVRSACPFNKGISQ 377
QY      178 -DIDRNNTPMWS--RPHNPMVYLPIFMHGKPKRSP-----NPSHEAKOPTPNEF 225
DB      378 GDLKTDITPLAGVYQPHGQKDVSNFAMGSPILKPGPETLCEFSSPCSQGARFLATEH 437
QY      226 RAPELK----FOVSVYKKAEDLMGTSDLMFGYTOQSHMQJFNGKNSRPFVHDYQPE 280
DB      438 QERGLMGDKMYNGVRPOIK-CEQSYGTGNSDESG-----SFSEADSESCPVQDRGQE 488
QY      281 IFLTPQV--YSGLPMDGKVMIGMAVHSHNGESAKL-----SRSMNR 321
DB      489 VKLPFPVDQITDLPKRDPMIMK---HKLTSEQLFEFHIDIRRSRKNR 533

RESULT 13
YL70_ARCFU STANDARD; PRT; 324 AA.
AC 028112;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, last sequence update)
DT 20-AUG-2001 (Rel. 40, last annotation update)
DE HYPOPHETICAL PROTEIN AF2170.
GN AF2170.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,

```

```

RA      Ketchum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
RA      Richardson D.L., Kerlavage A.R., Graham D.E., Kyriades N.C.,
RA      Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA      Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA      Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA      Overbeek R., Cocayne J.D., Meldrum J.F., McDonald L., Utterback T.,
RA      Cotton M.D., Springs T., Arlanch P., Kaine B.P., Sykes S.M.,
RA      Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA      Mason T.M., Olsen G.J., Fraser C.M., Smith R.O., Woese C.R.,
RA      Venter J.C.;
RT      "The complete genome sequence of the hyperthermophilic, sulphate-
RT      reducing archaeon Archaeoglobus fulgidus.";
RL      Nature 390:364-370(1997).
CC      - SIMILARITY: CONTAINS 2 KELCH REPEATS.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC      or send an email to license@lsb.slb.ch).
DR      EMBL; AE000955; AAB89093.1; -.
DR      TIGR; AF2170; -.
DR      InterPro; IPR001798; Kelch.
DR      Pfam; PF01344; Kelch; 2.
KW      Hypothetical protein; Repeat; Complete proteome.
FT      REPEAT 229 276 KELCH 1.
FT      REPEAT 277 323 KELCH 2.
SQ      SEQUENCE 324 AA; 36025 MW; C44538CBFE96E45 CRC64;

Query Match
Best local Similarity 20.1%; Score 92; DB 1; Length 324;
Matches 54; Conservative 44; Mismatches 72; Indels 98; Gaps 18;

QY      224 EFRAPELKFOYSVKVKAED-----LMCTSDSLMFGYTOQSHMQJF-NGKNSRFRFH 275
DB      63 EFRPPN-RLEI---LTSSEDFSGREAAVMDGEIILIFGTF---VFENKYSPTDILL 112
QY      276 DYQPEIFLQPVYSDLP-----W-DGKVMI-----CGMAVHSHNGESAKLSRS 318
DB      113 SFNPKLERLVNLASLPHPISDYAAVWGDSTRVYIFLNSRCEVYAFYPSNESFALDVS 172
QY      319 -----W--NRAVL-----MAGMEMKNTLVMPRINGRIPEKSGSQ 351
DB      173 CPLEHGGCVHVVWYGKAYFCGEGVASFEDPMGFRW--IAFTDRAVWVRAATVADG-- 228
QY      352 PDDNPDLIDYGYGDAVRELYOLENKSNISET---VRIYPRSGKALOLDYVPLRGK-- 405
DB      229 -----YIFAIGSSGIATKDEILIRFNPKTGE-LCEMRTKLPVARGQA 270
QY      406 --ISG-YFOIF-----QGYQSLIDYNH 425
DB      271 VAVGERTYIFGGYTKDGTGIANETLRTDY 298

RESULT 14
RI14_HUMAN STANDARD; PRT; 1158 AA.
AC P48352;
DT 01-FEB-1996 (Rel. 33, Created)
DT 20-AUG-2001 (Rel. 40, last sequence update)
DT 20-AUG-2001 (Rel. 40, last annotation update)
DE NUCLEAR FACTOR RIP140 (NUCLEAR RECEPTOR INTERACTING PROTEIN 1).
GN NR1P1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```


RC TISSUE-Breast: PubMed-7641693;
RX MEDLINE-95369246;
RA Cavailles V., Dauvois S., L'Horsset F., Lopez G., Hoare S.,
RA Kusner P.J., Parker M.G.;
RT "Nuclear factor R1P140 modulates transcriptional activation by the
RT estrogen receptor";
RL EMO J. 14:3741-3751(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-20289799; PubMed-10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishi K., Tokoi Y., Choi D.-K., Soeda E.,
RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Norisaki G., Hornischer K., Brandt P.,
RA Scharte M., Schoen O., Desario A., Reichelt J., Kauer G., Bioecker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,
RA Weinmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehnach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
CC -1- FUNCTION: MODULATES TRANSCRIPTIONAL ACTIVATION BY THE ESTROGEN
CC RECEPTOR.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X84373; CAAS9108.1; -
DR EMBL: AL163207; CAB90396.1; -
DR MIM: 602490; -
KM Transcription regulation; Nuclear protein.
FT CONFLICT 124 124 P->R (IN REF. 1).
FT CONFLICT 448 448 R->G (IN REF. 1).
FT CONFLICT 721 726 NKGSE -> TKGRVK (IN REF. 1).
FT CONFLICT 1080 1080 T -> A (IN REF. 1).
SQ SEQUENCE 1158 AA; 126941 MW; 81FC424968B9A5F6 CRC64;

Query Match 3.9%; Score 92; DB 1; Length 1158;
Best Local Similarity 20.0%; Pred. No. 21;
Matches 93; Conservative 62; Mismatches 138; Indels 172; Gaps 23;

OY 41 NDLOGDNELPIDVGSATOSASIDTANPLDEHEPELYTTA-----LENKTYMLIN----- 88
DB 524 NDVSFNFQNVARTSVISLSPSTNRTTPVS--TPPLTSSKAGSPINLSOSHSLVITWNSPP 581
OY 89 --CSALNDIRKLACYDTL-----VHGEPYAVIK-----TKRSIRLDETIMOTIKG 133
DB 582 YVCSQSQSKLNTASNSHMDLTKSKDPPEKPAONEGAMNSNTFSKILLOMLAOCQMOS 641
OY 134 POVYQOETTPDIFLNGNEKGMJTKDAKOLEYAAKQFTPLSLSPDLDRNNPPL----- 186
DB 642 SMSVEGRPSKQLTGN-----TDK-----PIGM-----IDRLNSPLSNKNTNA 680
OY 187 -----WSSRPINP-----MYVLPFPHGKPRKSRNTPSHKQKTP--- 222
DB 681 VEENKAFFSQPGPEPGSGSEIENLERRTYLQLLT-GNPKGKS-----EKKEETPLRD 735
OY 223 -----NEFRAPELKFOYSKVYKAEADLWGTDSDLMFGVTQOSHQJNGKNSRPF----- 272
DB 736 ESTGHSERALSQILMKYIKISEPCDDLOIPNTV-----HLSHDAKSAPFLGMA 785
OY 273 -RVHDYDQ-----ELFLTPQVYSDLPMDGKVRMIGMGAHVHSHNGESAKLSRSWNRAYL-- 324
DB 324

DB 786 PAYORSAPALPYSEDFKSEPV---SPQDFS-----FSKNGLLSLRLQNDSTLYAD 833
OY 325 -----WAGMKWKLTYMPRIWGRIFKEGSSQDDNP-----DILDYIG----- 363
DB 834 DSDSRHNNENALLSKNLCVWP-----KRKLTEPLENPFKMKKNIVDAANNHAP 887
OY 364 --YGDVRFVLEENKNSNSGVRNPRSGKALQLDVYVPLGK 405
DB 888 EYLSGLNLOEBLFRN-----DLERKYPAGHG 916

RESULT 15
CATA.DICDI STANDARD: PRT: 496 AA.
ID CATA.DICDI
AC 077229;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CATALASE (EC 1.11.1.6).
GN CATA OR CAT.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RA Foote C., Alexander H., Alexander S.;
RL Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE.
CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -1- COFACTOR: HEME GROUP (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF090443; AAC36743.1; -
DR HSSP: P00432; 7CAT.
DR DictyDb: DD07777; cata.
DR InterPro: IPR002226; Catalase.
DR Pfam: PF00199; catalase.1.
DR PRINTS: PR00067; CATALASE.
DR PRODOM: PD000510; Catalase.1.
DR PROSITE: PS00437; CATALASE_1; 1.
DR PROSITE: PS00438; CATALASE_2; 1.
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Peroxisome.
FT ACT_SITE 54 54 BY SIMILARITY.
FT ACT_SITE 128 128 BY SIMILARITY.
FT BINDING 338 338 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT SITE 494 496 MICROBODY TARGETING SIGNAL (POTENTIAL).
SQ SEQUENCE 496 AA; 55683 MW; 683318B8FBAED2E6 CRC64;

Query Match 3.9%; Score 91; DB 1; Length 496;
Best Local Similarity 23.5%; Pred. No. 7.9;
Matches 56; Conservative 32; Mismatches 94; Indels 56; Gaps 12;

OY 61 STDTPANPLDEHEPELYTTALENKTYMLINCSALNODIMRLACYDTLVHGETPAVVIKTRSI 120
DB 7 TTSSGSPIDNN-----LNSMTAGVNGPILIDFTLI-----DKLAHNDRE----- 46
OY 121 RLDETIMOTIKGPVYQOETTD-----PIF-----LMGNEKGMJ-T 156
DB 47 RPERIVVAH-KGAGAHGFEVTSVDVPRKAKAKFLNKVGKRPITPTSTVSGEGKSSDS 105

QY 157 KDKAQLEAYAKQFTPLSLSEFDLRNNTPLUMSSRPINPMYVLPITFMHGKPNRSPNTPSHE 216
DB 106 ERDPR-GFAVKFYTEEG-NEDWVGNNTPVPEFLRDPK---FPDFLHTQ-KRNPQTCKD 158
QY 217 AKQFTPNFRAPELKFOVSVKVKAEDLMGTDSDLWFGYTOOSHMOIFNGKNSRPFV 274
DB 159 PNMFMDFLGOTPESTHQVSLFSDR---GTPKSYRHHMGFSSTHLKFNVAQCKPYW 212

Search completed: November 30, 2001, 14:27:06
Job time: 553 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 13:05:18 ; Search time 77.39 Seconds
(without alignments)
835.410 Million cell updates/sec

Title: US-09-787-083-8
Perfect score: 442
Sequence: 1 MKVSLSTFLSLPCFALLA.....YHKAATSPGVGLMDNMGL 442

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 473505 seqs, 146272329 residues

Word size : 0

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPREMBL.17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	2.0	382	2 09K0U7	09K0U7 neisseria m
2	9	2.0	409	2 09J7J21	09J7J21 neisseria m
3	9	2.0	862	2 09CL62	09CL62 pasteurella
4	9	2.0	1432	10 09AX85	09AX85 oryza sativ
5	8	1.8	299	10 09M268	09M268 arabidopsis
6	8	1.8	306	2 09CL22	09CL22 pasteurella
7	7	1.6	55	12 067608	067608 tomato gold
8	7	1.6	72	12 0912S5	0912S5 potato viru
9	7	1.6	72	12 0912S4	0912S4 potato viru
10	7	1.6	95	12 0912S3	0912S3 potato viru
11	7	1.6	96	12 0912S8	0912S8 potato viru
12	7	1.6	101	12 0912S6	0912S6 potato viru
13	7	1.6	122	2 09RU16	09RU16 delnoccoc
14	7	1.6	124	12 0912S7	0912S7 potato viru
15	7	1.6	131	6 09GMD4	09GMD4 macaca fusc
16	7	1.6	138	12 0912Q7	0912Q7 potato viru
17	7	1.6	147	5 036035	036035 hexamita in
18	7	1.6	158	2 0929C1	0929C1 chlamydia p
19	7	1.6	169	2 0912A7	0912A7 streptomyc

20	7	1.6	186	10 09LRX9	09LRX9 arabidopsis
21	7	1.6	189	13 09DFT2	09DFT2 notochenia
22	7	1.6	198	4 09BUX9	09BUX9 homo sapien
23	7	1.6	212	13 091819	091819 ictalurus p
24	7	1.6	214	5 09VTR5	09VTR5 drosophila
25	7	1.6	216	2 09X6H6	09X6H6 streptococ
26	7	1.6	218	2 09RFL4	09RFL4 streptococ
27	7	1.6	219	5 017594	017594 caenorhabd
28	7	1.6	220	2 09XCK1	09XCK1 streptococ
29	7	1.6	220	5 036036	036036 spironucleu
30	7	1.6	221	12 065002	065002 alfalfa mos
31	7	1.6	227	5 09BME4	09BME4 glosina mo
32	7	1.6	233	6 09MZB0	09MZB0 ovis arles
33	7	1.6	236	13 09DPS8	09DPS8 notochenia
34	7	1.6	264	12 085260	085260 potato viru
35	7	1.6	265	9 038155	038155 bacterioph
36	7	1.6	267	9 099363	099363 bacterioph
37	7	1.6	267	12 085261	085261 potato viru
38	7	1.6	267	12 098628	098628 potato viru
39	7	1.6	267	12 086842	086842 potato viru
40	7	1.6	267	12 099026	099026 potato viru
41	7	1.6	267	12 005476	005476 potato viru
42	7	1.6	267	12 009711	009711 potato viru
43	7	1.6	267	12 09WH17	09WH17 potato viru
44	7	1.6	267	12 09QNM9	09QNM9 potato viru
45	7	1.6	267	12 09DUL0	09DUL0 potato viru

ALIGNMENTS

RESULT 1
ID 09K0U7 PRELIMINARY; PRT; 382 AA.
AC 09K0U7;
DT 01-OCR-2000 (TREMREL. 15, Created)
DT 01-OCR-2000 (TREMREL. 15, Last sequence update)
DT 01-JUN-2001 (TREMREL. 17, Last annotation update)
DE PHOSPHOLIPASE A1, PUTATIVE.
GN NMB0464.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J., Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A., Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B., Cotton M.D., Uitterlinden T.R., Khouri H., Qin H., Yamathavan J., Gill J., Scarlato V., Masignani V., Piza M., Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappelli R., Venter J.C.;
RA "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.";
RT Science 287:1809-1815(2000).
RL EMBL: AE002403; AAF40901.1; -;
DR TIGR: NMB0464; -;
DR InterPro: IPR003187; PLAI.
DR Pfam: PF02253; PLAI; 1.
KW Complete proteome.
SQ SEQUENCE 382 AA; 42714 MW; B468A802F062E836 CRC64;

Query Match 2.0%; Score 9; DB 2; Length 382;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 279 PEIRTPQPV 287
|||||||
Db 221 PEIRTPQPV 229

```
RESULT 2
Q9J7J1 PRELIMINARY: PRT: 409 AA.
AC 09J7J1:
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PUTATIVE PHOPHOLIPASE.
GN NMA2021.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=63699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Felwell T., Hamlin N., Holtroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491."
RL Nature 404:502-506(2000).
DR EMBL: AL162757; CAB85240.1;
DR InterPro: IPR003187; Pfam:
DR Pfam: PR02253; Pfam: 1.
KW Complete proteome.
SQ SEQUENCE 409 AA; 45862 MW; CD6585B064D01A41 CRC64;

Query Match
Best local Similarity 100.0%; Score 9; DB 2; Length 409;
Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIPLTQPV 287
DB 248 PEIPLTQPV 256

RESULT 3
Q9CL62 PRELIMINARY: PRT: 862 AA.
AC 09CL62:
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CVGS1.
GN CVGSY OR PM1380.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang O., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida pm70."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.
CC -1- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC KINASES.
DR EMBL: AE006176; AA033464.1;
DR InterPro: IPR000410; BcTL_sensory.
DR InterPro: IPR000658; DUF5.
DR InterPro: IPR003660; HAMF.
DR InterPro: IPR003594; HATPase_C.
DR InterPro: IPR003661; His_kinA.

```

```
DR InterPro: IPR001789; Response_reg.
DR Pfam: PF00672; DUF5. 1.
DR Pfam: PF00072; response_reg. 1.
DR Pfam: PF00512; signal. 1.
DR PRINTS: PR00344; BCTRLSENSOR.
DR SMART: SM00304; HAMF. 1.
DR SMART: SM00387; HATPase_C. 1.
DR SMART: SM00388; HSKA. 1.
DR SMART: SM00448; RBC. 1.
KW Complete proteome; Kinase; Phosphorylation; Sensory transduction;
KW Transferrase.
SQ SEQUENCE 862 AA; 97846 MW; AB6814A3B63626C0 CRC64;

Query Match
Best local Similarity 100.0%; Score 9; DB 2; Length 862;
Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LSTPLTSL 13
DB 293 LSTPLTSL 301

RESULT 4
Q9AX85 PRELIMINARY: PRT: 1432 AA.
AC 09AX85:
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PUTATIVE ABC TRANSPORTER PROTEIN.
GN P0410E03.7.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone: P0410E03."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP002844; BAB21276.1;
SQ SEQUENCE 1432 AA; 160234 MW; 04E8CF2CA08D993 CRC64;

Query Match
Best local Similarity 100.0%; Score 9; DB 10; Length 1432;
Pred. No. 4.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 FTPLSLFSD 178
DB 830 FTPLSLFSD 838

RESULT 5
Q9M268 PRELIMINARY: PRT: 299 AA.
AC 09M268:
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE RAV-LIKE PROTEIN.
GN F21F14.140.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.

```

RA Choisine N., Robert C., Brothier P., Wincker P., Catolico L.,
 RA Attigeneve F., Saurin M., Weissenbach J., Mewes H.W., Lemcke K.,
 RA Mayer K.F.X., Queller F., Salanoubat M.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL38642; CAB/1904.1; -;
 DR InterPro: IPR003340; B3.
 DR Pfam: PF02362; B3; 1.
 SQ SEQUENCE 299 AA; 34270 MW; 2AB841F179DD174B CRC64;

Query Match 1.8%; Score 8; DB 10; Length 299;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 SLSPDLR 181
 DB 292 SLSPDLR 299

RESULT 6
 ID 09CL22 PRELIMINARY; PRT; 306 AA.
 AC 09CL22;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE HYPOTHETICAL PROTEIN PM1426.
 GN PM1426.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Pasteurella.
 OX NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PM70;
 RX MEDLINE-21145866; PubMed-11248100;
 RA May B.J., Zhang O., Li L., Paustian M.L., Whittam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida pm70.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 DR EMBL: AB006179; AAK03510.1; -;
 DR InterPro: IPR003187; PLAI.
 DR Pfam: PF02253; PLAI; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 306 AA; 35580 MW; EAF3DE8C1C22B26E CRC64;

Query Match 1.8%; Score 8; DB 2; Length 306;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 314 KLSRSMR 321
 DB 184 KLSRSMR 191

RESULT 7
 ID 067608 PRELIMINARY; PRT; 55 AA.
 AC 067608;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE SUBGENOMIC DNA DERIVED FROM DNA B CCCDS - COVALENTLY CLOSED CIRCULAR
 DE DOUBLE-STRANDED MOLECULE.
 OS Tomato golden mosaic virus (TGMV).
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
 OX NCBI_TaxID=10831;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE-87040767; PubMed-3022243;
 RA MacDowell S.W., Coutts R.H.A., Buck K.W.;
 RT "Molecular characterization of subgenomic single-stranded and double-
 RT stranded DNA forms isolated from plants infected with tomato golden
 RT mosaic virus.";
 RL Nucleic Acids Res. 14:7967-7984(1986).
 DR EMBL: X04485; CAA28171.1; -;
 DR InterPro: IPR000211; Gemin1.BL.
 DR Pfam: PF00845; Gemin1.BL; 1.
 SQ SEQUENCE 55 AA; 6281 MW; 41963570D739EC96 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 55;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 312 SAKLSRS 318
 DB 43 SAKLSRS 49

RESULT 8
 ID 091ZS5 PRELIMINARY; PRT; 72 AA.
 AC 091ZS5;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE COAT PROTEIN (FRAGMENT).
 GN CP.
 OS Potato virus Y strain NTN (PVY (NTN)).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Potyvirus.
 OX NCBI_TaxID=122280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TUBER NECROSING (NTN);
 RA Cerovska N., Filigaro M., Moravec T., Petrzik K.;
 RT "Differences in nucleotide and amino acid sequences of N-terminal
 RT parts of coat proteins among isolates of potato virus Y NTN strain.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF228633; AAF67874.1; -;
 FT NON_TER 1 72
 FT NON_TER 1 72
 SQ SEQUENCE 72 AA; 7721 MW; 2C5287E15624CBDD CRC64;

Query Match 1.6%; Score 7; DB 12; Length 72;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 TKDKAQ 162
 DB 11 TKDKAQ 17

RESULT 9
 ID 091ZS4 PRELIMINARY; PRT; 72 AA.
 AC 091ZS4;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE COAT PROTEIN (FRAGMENT).
 GN CP.
 OS Potato virus Y strain NTN (PVY (NTN)).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Potyvirus.
 OX NCBI_TaxID=122280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TUBER NECROSING (NTN);
 RA Cerovska N., Filigaro M., Moravec T., Petrzik K.;

RT "Differences in nucleotide and amino acid sequences of N-terminal
 parts of coat proteins among isolates of potato virus Y-NTN strain.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF228634; AAF67875.1; -.
 FT NON_TER 1 1
 FT NON_TER 72 72
 SQ SEQUENCE 72 AA; 7692 MW; 4F129B870523B482 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 72;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAQ 162
 Db 11 TKKDAQ 17

RESULT 10
 ID Q91ZS3 PRELIMINARY; PRT; 95 AA.
 AC Q91ZS3;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DE COAT PROTEIN (FRAGMENT).
 GN CP.
 OS Potato virus Y strain NTN (PVY(NTN)).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Polyvirus.
 OX NCBI_TaxID=122280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TUBER NECROSING (NTN);
 RA Cerovska N., Filiigayova M., Moravec T., Petrizik K.;
 RT "Differences in nucleotide and amino acid sequences of N-terminal
 parts of coat proteins among isolates of potato virus Y-NTN strain.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF228635; AAF67876.1; -.
 FT NON_TER 1 1
 FT NON_TER 95 95
 SQ SEQUENCE 95 AA; 10446 MW; 28C06891777C9928 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 95;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAQ 162
 Db 11 TKKDAQ 17

RESULT 11
 ID Q91ZS8 PRELIMINARY; PRT; 96 AA.
 AC Q91ZS8;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE COAT PROTEIN (FRAGMENT).
 GN CP.
 OS Potato virus Y (strain N) (PVY).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Polyvirus.
 OX NCBI_TaxID=12219;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NECROSING (N);
 RA Cerovska N., Filiigayova M., Moravec T., Petrizik K.;
 RT "Differences in nucleotide and amino acid sequences of N-terminal
 parts of coat proteins among isolates of potato virus Y-NTN strain.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF228630; AAF67871.1; -.
 FT NON_TER 1 1
 FT NON_TER 96 96
 SQ SEQUENCE 96 AA; 10559 MW; 89C3A5675B9784CA CRC64;

Query Match 1.6%; Score 7; DB 12; Length 96;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAQ 162
 Db 11 TKKDAQ 17

RESULT 12
 ID Q91ZS6 PRELIMINARY; PRT; 101 AA.
 AC Q91ZS6;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE COAT PROTEIN (FRAGMENT).
 GN CP.
 OS Potato virus Y strain NTN (PVY(NTN)).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Polyvirus.
 OX NCBI_TaxID=122280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TUBER NECROSING (NTN);
 RA Cerovska N., Filiigayova M., Moravec T., Petrizik K.;
 RT "Differences in nucleotide and amino acid sequences of N-terminal
 parts of coat proteins among isolates of potato virus Y-NTN strain.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF228632; AAF67873.1; -.
 FT NON_TER 1 1
 FT NON_TER 101 101
 SQ SEQUENCE 101 AA; 11112 MW; B55BF51C3D5FE9BD CRC64;

Query Match 1.6%; Score 7; DB 12; Length 101;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAQ 162
 Db 11 TKKDAQ 17

RESULT 13
 ID Q9RU16 PRELIMINARY; PRT; 122 AA.
 AC Q9RU16;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE HYPOTHETICAL 13.7 KDA PROTEIN.
 GN DR1580.
 OS Deinococcus radiodurans.
 OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
 OX NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RL;
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Yamakawa J.J., Lam P., McDonald L., Uitterback T., Zalewski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;

RT "genome sequence of the radioresistant bacterium Deinococcus
 RT radiodurans RI.";
 RL Science 286:1571-1577(1999).
 DR EMBL: AE002002; AAF1150.1; -.
 DR TIGR: DR1580; -.
 DR InterPro: IPR000325; Glyoxalase_1.
 DR Pfam: PF00903; Glyoxalase: 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 122 AA; 13683 MW; 7BA836DCAD786E51 CRC64;

Query Match 1.6%; Score 7; DB 2; Length 122;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 EHEPELY 76
 |||||
 Db 115 EHEPELY 121

RESULT 14
 Q91ZS7 PRELIMINARY; PRT; 124 AA.
 AC Q91ZS7;
 DT 01-OCT-2000 (TREMUREL. 15, Created)
 DT 01-OCT-2000 (TREMUREL. 15, Last sequence update)
 DT 01-OCT-2000 (TREMUREL. 15, Last annotation update)
 DE COAT PROTEIN (FRAGMENT).
 GN CP.
 OS Potato virus Y strain NTN (PVY(NTN)).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Polyviruses.
 OX NCBI_TaxID-122280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TUBER NECROSING (NTN):
 RA Cerowska N., Filigorska M., Moravec T., Petrzik K.;
 RT "Differences in nucleotide and amino acid sequences of N-terminal
 parts of coat proteins among isolates of potato virus Y-NTN strain."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF228631; AAF67872.1; -.
 FT NON_TER 1
 FT NON_TER 124
 SQ SEQUENCE 124 AA; 13654 MW; DE3618D620BA7080 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 124;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAQ 162
 |||||
 Db 11 TKKDAQ 17

RESULT 15
 Q9GMD4 PRELIMINARY; PRT; 131 AA.
 AC Q9GMD4;
 DT 01-MAR-2001 (TREMUREL. 16, Created)
 DT 01-MAR-2001 (TREMUREL. 16, Last sequence update)
 DT 01-JUN-2001 (TREMUREL. 17, Last annotation update)
 DE ALPHA-TUBULIN (FRAGMENT).
 OS Macaca fuscata (Japanese macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID-9542;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99270338; PubMed=10340513;
 RA Okuno H., Tokuyama W., Li Y.X., Hashimoto T., Miyashita Y.;
 RT "Quantitative evaluation of neurotrophin and trk mRNA expression in

RT visual and limbic areas along the occipito-temporo-hippocampal pathway
 RT in adult macaque monkeys.";
 RL J. Comp. Neurol. 408:378-398(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Hashimoto T., Tokuyama W., Okuno H., Li Y., Miyashita Y.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF286378; AAC00533.1; -.
 DR InterPro: IPR003008; Tubulin_FtsZ.
 DR Pfam: PF00091; tubulin; 1.
 KW GTP-binding.
 FT NON_TER 1
 FT NON_TER 131
 SQ SEQUENCE 131 AA; 14575 MW; A07D9C90B0641RCE CRC64;

Query Match 1.6%; Score 7; DB 6; Length 131;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 IKTKRSI 120
 |||||
 Db 110 IKTKRSI 116

Search completed: May 9, 2002, 13:05:19
 Job time: 614 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 13:05:53 ; Search time 28.06 Seconds
(without alignments)
577.543 Million cell updates/sec

Title: US-09-787-083-8
Perfect score: 442
Sequence: 1 MKVSLSTLTSLPCFALIA.....YNHEATSPGVGLMLNDMMGL 442

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 100059 seqs, 36664827 residues

Word size : 0

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	1.8	286	1	PAL_KLEPN
2	8	1.8	1396	1	PLTF_BPPT5
3	7	1.6	186	1	PRL_ICTPU
4	7	1.6	197	1	TBA_PLAYO
5	7	1.6	240	1	TBA_OCTVU
6	7	1.6	270	1	Y304_MYCPN
7	7	1.6	273	1	ILVE_THEMA
8	7	1.6	322	1	MIAA_PSEPU
9	7	1.6	326	1	SPL_BACBR
10	7	1.6	342	1	FLAD_METJA
11	7	1.6	342	1	Y735_METJA
12	7	1.6	423	1	AMT_AQUAE
13	7	1.6	444	1	TBA_ONCKE
14	7	1.6	448	1	TBA4_HUMAN
15	7	1.6	448	1	TBA5_CHICK
16	7	1.6	449	1	TBA2_DROME
17	7	1.6	449	1	TBAE_PHYPO
18	7	1.6	449	1	TBAN_PHYPO
19	7	1.6	449	1	TBA_PHYPO
20	7	1.6	450	1	TBAD_PHYPO
21	7	1.6	450	1	TBA_HARCO
22	7	1.6	450	1	TBA_NORVI
23	7	1.6	451	1	TBA1_CRIGR
24	7	1.6	451	1	TBA3_HOMAM
25	7	1.6	451	1	TBA_OCTDO
26	7	1.6	451	1	TBA_TORMA
27	7	1.6	511	1	LGT_CITUN
28	7	1.6	511	1	ATPA_HAEIN
29	7	1.6	536	1	60IM_BUCAP
30	7	1.6	544	1	GPV_RAT
31	7	1.6	567	1	GPV_RAT
32	7	1.6	610	1	FIMB_DICDI
33	7	1.6	635	1	DNAM_HALMA

34	7	1.6	768	1	YB23_HUMAN	Q9u1j7	homo sapien
35	7	1.6	783	1	FYB_HUMAN	O15117	homo sapien
36	7	1.6	900	1	AXN1_HUMAN	O15169	homo sapien
37	7	1.6	947	1	LKTA_PASBP	P55123	pasteurella
38	7	1.6	1026	1	VG37_BPPT4	P03744	bacterioph
39	7	1.6	3061	1	POLG_PYTHU	O02963	p genome po
40	6	1.4	33	1	BR2B_RANES	P40838	rana escul
41	6	1.4	33	1	BR2E_RANES	P32413	rana escul
42	6	1.4	50	1	YK05_CABEL	P34300	caenorhabd
43	6	1.4	51	1	INS_ACOCA	P01324	acomys cahl
44	6	1.4	51	1	INS_ALIMI	P12703	alligator m
45	6	1.4	51	1	INS_ANSAN	P07454	anser anser

ALIGNMENTS

RESULT 1	ID	STANDARD	PRT	286 AA.
PAL_KLEPN	AC	P37446		
01-OCT-1994 (Rel. 30, Created)	DT			
01-OCT-1994 (Rel. 30, Last sequence update)	DT			
01-NOV-1997 (Rel. 35, Last annotation update)	DT			
PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DEMERGENT-RESISTANT	DE			
PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-	DE			
ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).	GN			
PLDA.	OS			
Klebsiella pneumoniae.	OC			
Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;	OC			
Klebsiella.	OX			
NCBI_TaxID=573;	OX			
[1]	RN			
SEQUENCE FROM N.A.	RP			
MEDLINE=94131966; PubMed=8300539;	RX			
Brok R.G.P.M., Brinkman E., van Boxtel R., Bekkers A.C.A.P.,	RA			
Verheij H.M., Tomassen J.;	RT			
Molecular characterization of enterobacterial plda genes encoding	RT			
outer membrane phospholipase A.;	RL			
Bacteriol. 176:861-870(1994).	CC			
- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE	CC			
A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.	CC			
- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O -> 1-ACYLGLYCERO-	CC			
PHOSPHOCHOLINE + A FATTY ACID ANION.	CC			
- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O -> 2-ACYLGLYCERO-	CC			
PHOSPHOCHOLINE + A FATTY ACID ANION.	CC			
- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.	CC			
- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES	CC			
LOCATED THERE.	CC			
-----	CC			
This SWISS-PROT entry is copyright. It is produced through a collaboration	CC			
between the Swiss Institute of Bioinformatics and the EMBL outstation -	CC			
the European Bioinformatics Institute. There are no restrictions on its	CC			
use by non-profit institutions as long as its content is in no way	CC			
modified and this statement is not removed. Usage by and for commercial	CC			
entities requires a license agreement. (See http://www.isb-sib.ch/announce/	CC			
or send an email to license@isb-sib.ch .)	CC			
-----	CC			
EMBL: X76901; CAA54223.1; -	DR			
PIR: B36971; B36971.	DR			
PIR: S40129; S40129.	DR			
InterPro: IPR003187; PLA1.	DR			
Pfam: PF02253; PLA1; 1.	KW			
Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium.	FT			
SIGNAL	FT			
CHAIN 1 286	FT			
ACT SITE 161 161	FT			
BY SIMILARITY.	FT			
SEQUENCE 286 AA; 32544 MW; 3E39F863085108A3 CRC64;	SQ			

Query Match 1.8%; Score 8; DB 1; Length 286;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 431 GVGMLND 438
 DB 277 GVGMLND 284

RESULT 2

VFLE_BPT5 STANDARD; PRT: 1396 AA.
 AC P13390; 048502;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE L-SHAPED TAIL FIBER PROTEIN (LTF PROTEIN).
 GN LTF.
 OS Bacteriophage T5.
 OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae.
 OX NCBI_TaxID=10726;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-95309401; PubMed-7789514;
 RA Kallman A.V., Kulshln V.E., Shlyapnikov M.G., Ksenzenko V.N.,
 RA Kiyukov V.M.;
 RT "The nucleotide sequence of the bacteriophage T5 ltf gene.";
 RL FEBS Lett. 366:46-48(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kallman A.V.;
 RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
 RX MEDLINE-88289370; PubMed-3267228;
 RA Kallman A.V., Kiyukov V.M., Bayev A.A.;
 RT "The nucleotide sequence of bacteriophage T5 DNA at the region
 RT between early and late genes.";
 RL Nucleic Acids Res. 16:6230-6230(1988).
 CC -!- FUNCTION: NONESSENTIAL PROTEIN THAT MEDIATES BINDING TO THE
 CC POLYMANNOSE O ANTIGEN.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL: X69460; CAA49220.1; -;
 DR EMBL: AJ001191; CAA04591.1; -;
 DR PIR: S01982; S01982.
 KW Late protein.
 FT CONFICIT 986 986 V -> A (IN REF. 2).
 SQ SEQUENCE 1396 AA; 147989 MW; 18CD2192F65FFC1 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 1396;
 Best Local Similarity 100.0%; Pred. No. 6.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 54 QSATOSAS 61
 DB 87 QSATOSAS 94

RESULT 3

PRL_ICTPU STANDARD; PRT: 186 AA.
 AC P31504;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE PROLACTIN (PRL).
 OS Ictalurus punctatus (Channel catfish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Siluriformes; Ictaluridae; Ictalurus.
 OX NCBI_TaxID=7998;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Plutitary;
 RX MEDLINE-93364578; PubMed-1308206;
 RA Watanabe K., Igarashi A., Noso T., Chen T.T., Dunham R.A.,
 RA Kawachi H.;
 RT "Chemical identification of catfish growth hormone and prolactin.";
 RL Mol. Mar. Biol. Biotechnol. 1:239-249(1992).
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
 DR HSP: 028632; IAN3.
 DR InterPro: IPR001400; SOMATOTROPIN.
 DR Pfam: PF00103; hormone; 1.
 DR PRINTS: PR00836; SOMATOTROPIN.
 DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
 DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
 KW Hormone; Plutitary.
 FT DISULFID 45 159 BY SIMILARITY.
 FT DISULFID 176 186 BY SIMILARITY.
 SQ SEQUENCE 186 AA; 20606 MW; 7BCCD89718F44E74 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 186;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 41 NDIGQDN 47
 DB 140 NDIGQDN 146

RESULT 4

TBA_PLAYO STANDARD; PRT: 197 AA.
 AC P12543;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE TUBULIN ALPHA CHAIN (FRAGMENT).
 OS Plasmodium berghei yoelli.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5862;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-89014607; PubMed-2459618;
 RA Akella R., Arasu P., Vaidya A.B.;
 RT "Molecular clones of alpha-tubulin genes of Plasmodium yoelli reveal
 RT an unusual feature of the carboxy terminus.";
 RL Mol. Biochem. Parasitol. 30:165-174(1988).
 CC -!- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
 CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
 CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
 CC -!- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
 CC -!- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL: M29816; AAA29779.1; -;
 DR PIR: A54506; A54506.
 DR InterPro: IPR000217; Tubulin.
 DR InterPro: IPR003008; Tubulin_ftsz.
 DR Pfam: PF00091; tubulin; 1.
 DR PROSITE: PS00227; TUBULIN; PARTIAL.

KW Microtubules: GTP-binding.
 FT NON_TER 1
 SQ SEQUENCE 197 AA; 22111 MW; 4DAAF199CC6CD319 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 197;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120
 |||||
 Db 82 IKTKRSI 88

RESULT 5
 TBA_OCTVU STANDARD; PRT; 240 AA.
 AC P24635;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE TUBULIN ALPHA CHAIN (FRAGMENT).
 OS Octopus vulgaris (Octopus).
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda.
 CC Invertebrata; Octopodidae; Octopus.
 OX NCBI_TaxID=6645;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue-Lens;
 RA Zinov'Eva R.D., Aleinikova K.S., Tomarev S.I.;
 RT "Isolation and structural characterization of cDNAs coding for alpha-tubulin of the octopus eye-lens."
 RL Dokl. Akad. Nauk SSSR 302:462-467(1988).
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X15845; CAA33844.1; -.
 DR PIR: A61544; A61544.
 DR InterPro: IPR000217; Tubulin.
 DR InterPro: IPR003008; Tubulin_FtsZ.
 DR Pfam: PF00091; Tubulin; 1.
 DR PROSITE: PS00227; TUBULIN; PARTIAL.
 KW Microtubules; GTP-binding.
 FT NON_TER 1
 FT SITE 1
 SQ SEQUENCE 240 AA; 26961 MW; 13BB3A1F740F2416 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 240;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120
 |||||
 Db 124 IKTKRSI 130

RESULT 6
 Y304_MYCPN STANDARD; PRT; 270 AA.
 ID Y304_MYCPN
 AC P75355;
 DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHEICAL ABC TRANSPORTER ATP-BINDING PROTEIN MG304 HOMOLOG (A05_0RF270L).
 DE (A05_0RF270L).
 GN MPN433 OR MP408.
 OS Mycoplasma pneumoniae.
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 CC Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RX MEDLINE=97105885; PubMed=8948633;
 RA Himmelfreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C., Herrmann R.;
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae."
 RL Nucleic Acids Res. 24:4420-4449(1996).
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: A600040; AAB96056.1; -.
 DR InterPro: IPR003593; AAA.
 DR InterPro: IPR003439; ABC_Transport.
 DR InterPro: IPR001687; ATP_GTP_A.
 DR Pfam: PF00005; ABC_tran; 1.
 DR SMART: SM00382; AAA; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 KW Hypothetical protein; ATP-binding; Transport; Complete proteome.
 FT NP_BIND 36
 FT SITE 43
 SQ SEQUENCE 270 AA; 30770 MW; 6512640B4BC051B4 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 270;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 168 KQTFPLS 174
 |||||
 Db 125 KQTFPLS 131

RESULT 7
 LIVE_THEMA STANDARD; PRT; 273 AA.
 ID LIVE_THEMA
 AC P74921;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PROBABLE BRANCHED-CHAIN AMINO ACID AMINOTRANSFERASE (EC 2.6.1.42) (BCAT).
 DE (BCAT).
 GN LIVE OR TM0831.
 OS Thermotoga maritima.
 CC Bacteria; Thermotogales; Thermotoga.
 OX NCBI_TaxID=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J., Hart D.R., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A., McDonald L., Utterback T.R., Malek J.A., Liner K.D., Garrett M.M., Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D., Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O., Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;

```

RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima."
RL Nature 399:323-329(1999).
RN (2)
RP SEQUENCE OF 68-273 FROM N.A.
RC STRAIN-MSB8 / DSM 3109;
RX MEDLINE=97017137; PubMed=8863738;
RA Guipaud O., Labeidan B., Forcette P.;
RT "A gyrB-like gene from the hyperthermophilic bacterium Thermotoga
RL maritima."
RL Gene 174:121-128(1996).
CC -1- CATALYTIC ACTIVITY: L-LEUCINE + 2-OXOGlutARATE -> 4-METHYL-2-
CC OXOBUTANOATE + L-GLUTAMATE (ALSO ACTS ON L-ISOLEUCINE AND
CC L-VALINE).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: VALINE AND ISOLEUCINE BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO CLASS-IV OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001750; AAC35913.1; -.
DR EMBL; U49692; AAC44497.1; -.
DR HSSP; P00510; 1A3G.
DR TIGR; TM0831; -.
DR InterPro; IPR001544; AminoTran_4.
DR Pfam; PF01063; aminoTran_4; 1.
DR ProDom; PD001961; AminoTran_4; 1.
DR PROSITE; PS00770; AA_TRANSFER_CLASS_4; 1.
KM Transferase; AminoTransferase; Branched-chain amino acid biosynthesis;
KM Pyridoxal phosphate; Complete proteome.
FT CONFLICT 77
SQ SEQUENCE 273 AA; 31158 MW; 2162B705612E90E3 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 273;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 PLSLSPD 178
Db 61 PLSLSFD 67
|||||
RESULT 8
MIAA_PSEPU STANDARD; PRT; 322 AA.
AC 030762;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE TRNA DELTA(2)-ISOPENTENYL-PYROPHOSPHATE TRANSFERASE (EC 2.5.1.8) (IPP
DE TRANSFERASE) (ISOPENTENYL-DIPHOSPHATE:TRNA ISOPENTENYLTRANSFERASE)
DE (IPTASE) (IPT).
GN MIAA.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=503;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M.
RA Olekhnovich I.N., Gusein G.N.;
RT "Attenuation of the Pseudomonas putida trpE and trpGDC genes.";
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
-1- FUNCTION: CATALYZES THE FIRST STEP IN THE BIOSYNTHESIS OF
2-METHYLTHIO-N6-(DELTA(2)-ISOPENTENYL)-ADENOSINE (MS[2]I[6]A1)

```

```

CC ADVACENT TO THE ANTICODON OF SEVERAL TRNA SPECIES (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ISOPENTENYL DIPHOSPHATE + TRNA -
CC PYROPHOSPHATE + TRNA CONTAINING 6-ISOPENTENYLADENOSINE.
CC -1- SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaborati
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercia
CC entities requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF016312; AAB69443.1; -.
DR InterPro; IPR002627; IPPP.
DR Pfam; PF01715; IPPP; 1.
DR ProDom; PD004674; IPPP; 1.
KM Transferase; Nucleotidyltransferase; trna processing; ATP-binding.
FT NP_BIND 12 19
SQ SEQUENCE 322 AA; 35488 MW; DIA8B1A2D7B32F6A CRC64;

Query Match 1.6%; Score 7; DB 1; Length 322;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 FVDEVRS 39
Db 231 FVDEVRS 237
|||||
RESULT 9
ID SPI_BACBR STANDARD; PRT; 326 AA.
SP1_BACBR
AC P43131;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PROTEASE INHIBITOR PRECURSOR (BBRP1).
DE Bacillus brevis.
OS Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Brevibacillus.
NCBI_TaxID=1393;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-32; 104-112 AND 122-136.
RC STRAIN=HPD31;
RX MEDLINE=92304060; PubMed=1610177;
RA Shiga Y., Hasegawa K., Tsuboi A., Yamagata H., Uda K.S.;
RT "Characterization of an extracellular protease inhibitor of Bacillus
RT brevis HPD31 and nucleotide sequence of the corresponding gene.";
RL Appl. Environ. Microbiol. 58:525-531(1992).
CC -1- FUNCTION: SHOWS INHIBITORY ACTIVITY TOWARDS SERINE PROTEASES, SUCH
CC AS TRYPSIN, CHYMOTRYPSIN, AND SUBSTITISIN. MAY FORM A TRYPSIN-
CC INHIBITOR COMPLEX IN A MOLAR RATIO OF 1:1. IT IS HEAT RESISTANT AT
CC NEUTRAL AND ACIDIC PHs.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- PTM: MAY UNDERGO ACTIVATION AFTER SECRETION. IT IS PRODUCED
CC EXTRACELLULARLY IN MULTIPLE FORMS HAVING AT LEAST THREE DIFFERENT
CC MOLECULAR WEIGHTS (BBRP1-A, -B, AND -C).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercia
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D10696; BAA01538.1; -.
DR Serine protease inhibitor; Protease inhibitor; Signal; Repeat.
KM SIGNAL 1 24
FT CHAIN 25 326 SERINE PROTEASE INHIBITOR.
FT CHAIN 104 326 SERINE PROTEASE INHIBITOR-C.

```

FT CHAIN 104 ? SERINE PROTEASE INHIBITOR-B (POTENTIAL).
 FT CHAIN 122 326
 FT DOMAIN 177 304 CONTAINS TWO APPROXIMATE REPEATS.
 FT REPEAT 177 208 1.
 FT REPEAT 272 304
 SQ SEQUENCE 326 AA; 35100 MM; LC0456ABFA912F77 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 326;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 30 PVAFVDE 36
 Db 119 PVAFVDE 125

RESULT 10
 FLAD_METUA STANDARD: PRT: 342 AA.
 AC Q58305;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PUTATIVE FLAGELLA-RELATED PROTEIN D.
 GN FLAD OR M0895.
 OS Methanococcus jannaschli.
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 OC Methanococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
 RA MEDLINE-96337999; PubMed-6688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Klenk H.-P., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Clontz H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschli.";
 RL Science 273:1058-1073(1996).
 CC -1- SIMILARITY: STRONG, TO M.VOLTAE FLAD, ALSO TO FLAE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U67533; AAB98898.1; -
 DR TIGR: M07055; -
 KW Flagella; Complete proteome.
 RM SEQUENCE 342 AA; 39950 MM; B384DDE175566C CRC64;
 SQ

Query Match 1.6%; Score 7; DB 1; Length 342;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 356 PDILDPY 362
 Db 255 PDILDPY 261
 RESULT 11
 Y755_METUA STANDARD: PRT: 342 AA.
 ID Y755_METUA

AC Q58165;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHETICAL PROTEIN M07055 PRECURSOR.
 GN M07055.
 OS Methanococcus jannaschli.
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 OC Methanococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
 RA MEDLINE-96337999; PubMed-6688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Clontz H.-P., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschli.";
 RL Science 273:1058-1073(1996).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U67521; AAB98748.1; -
 DR TIGR: M07055; -
 KW Hypothetical protein; Signal; Complete proteome.
 RM SEQUENCE 342 AA; 38778 MM; BD25220A/EEND85C CRC64;
 SQ

Query Match 1.6%; Score 7; DB 1; Length 342;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 76 YTTALEN 82
 Db 90 YTTALEN 96

RESULT 12
 AMT_AOUAE STANDARD: PRT: 423 AA.
 AC Q66515;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PROBABLE AMMONIUM TRANSPORTER.
 GN AMT OR AMTB OR AO_112.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VF5;
 RA MEDLINE-98196666; PubMed-9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aulay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus.";
 RL Nature 392:353-358(1998).

```

CC -1- FUNCTION: INVOLVED IN THE UPTAKE OF AMONIA (PROBABLE).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE AMT1/MEP/MRGA FAMILY OF AMMONIUM
CC TRANSPORTERS (TC 2.4.9).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE00674; AAC06478.1; -.
CC InterPro: IPR001905; Ammonium_transp.
CC Pfam: PF00909; Ammonium_transp. 1.
CC PROSITE: PS01219; AMMONIUM_TRANSPO. 1.
CC KM Transport; Transmembrane; Complete proteome.
CC FT TRANSMEM 3 23 POTENTIAL.
CC FT TRANSMEM 31 51 POTENTIAL.
CC FT TRANSMEM 64 84 POTENTIAL.
CC FT TRANSMEM 123 143 POTENTIAL.
CC FT TRANSMEM 150 170 POTENTIAL.
CC FT TRANSMEM 183 203 POTENTIAL.
CC FT TRANSMEM 217 227 POTENTIAL.
CC FT TRANSMEM 246 266 POTENTIAL.
CC FT TRANSMEM 278 298 POTENTIAL.
CC FT TRANSMEM 301 321 POTENTIAL.
CC FT TRANSMEM 331 351 POTENTIAL.
CC FT TRANSMEM 374 394 POTENTIAL.
CC SO SEQUENCE 423 AA; 44322 MW; 166203BBB2A8A15B CRC64;

Query Match 1.6%; Score 7; DB 1; Length 423;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 402 LKKGISG 408
DB 104 LKKGISG 110

RESULT 13
TBA_ONCKE STANDARD; PRT; 444 AA.
AC P30436;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE TUBULIN ALPHA CHAIN.
OS Oncorhynchus keta (Chum salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8018;
RN NCB1_TaxID=8018;
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93208376; PubMed=1296820;
RA Coe I.R., Munro R., Sherwood N.M.;
RT "Isolation of different brain-specific isoforms of alpha-tubulins
RT from chum salmon (Oncorhynchus keta).";
RL DNA Seq. 3:257-262(1992).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X66973; CAA47384.1; -.
CC PIR: S25004; S25004.
CC InterPro: IPR000217; Tubulin.
CC Pfam: PF003008; Tubulin_FrsZ.
CC PRINTS: PRO1091; tubulin; 1.
CC PRINTS: PRO1161; TUBULIN.
CC PRINTS: PRO1162; ALPHATUBULIN.
CC PROSITE: PS00227; TUBULIN; 1.
CC KM Microtubules; GTP-binding; Multigene family.
CC FT NP_BIND 136 142 GTP (POTENTIAL).
CC FT SITE 444 444 INVOLVED IN POLYMERIZATION.
CC SO SEQUENCE 444 AA; 49314 MW; C2CD75432DDDBAF CRC64;

Query Match 1.6%; Score 7; DB 1; Length 444;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120
DB 329 IKTKRSI 335

RESULT 14
TBA4_HUMAN STANDARD; PRT; 448 AA.
AC P05215;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE TUBULIN ALPHA-4 CHAIN.
GN TUBA4.
OS Homo sapiens (Human).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606, 9541, 10090;
RN NCB1_TaxID=9606, 9541, 10090;
RP SEQUENCE FROM N.A.
RC SPECIES=Human, and M. fascicularis;
RX MEDLINE=87146332; PubMed=3029670;
RA Dohner P.R., Kislauskis E., Wentworth B.M., Villa-Komaroff L.;
RT "Alternative 5' exons either provide or deny an initiator methionine
RT codon to the same alpha-tubulin coding region.";
RL Nucleic Acids Res. 15:199-218(1987).
RN NCB1_TaxID=9606, 9541, 10090;
RP SEQUENCE FROM N.A.
RC SPECIES=Mouse;
RX MEDLINE=87064538; PubMed=3785200;
RA Villante A., Wang D., Dohner P., Dolph P., Lewis S.A., Cowan N.J.;
RT "Six mouse alpha-tubulin mRNAs encode five distinct isoforms: testis-
RT specific expression of two sister genes.";
RL Mol. Cell. Biol. 6:2409-2419(1986).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
CC -1- MISCELLANEOUS: THIS TUBULIN DOES NOT HAVE A C-TERMINAL TYROSINE.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

DR EMBL: X06956; CAA30026.1; -;
 DR EMBL: X04757; CAA28453.1; -;
 DR EMBL: M13444; AAA40502.1; -;
 DR PIR: A25873; A25873.
 DR MGD: MGI:1095410; Tubu4.
 DR InterPro: IPR000217; Tubulin.
 DR InterPro: IPR003008; Tubulin_Ftsz.
 DR Pfam: PF00091; tubulin_1.
 DR PRINTS: PRO1161; TUBULIN.
 DR PRINTS: PRO1162; ALPHATUBULIN.
 DR PROSITE: PS00227; TUBULIN; 1.
 KM Microtubules; GTP-binding; Multigene family.
 FT NP_BIND 142 148 GTP (POTENTIAL).
 SQ SEQUENCE 448 AA; 49924 MW; C00ED90A183FE8F2 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 448;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 114 IKTKRSI 120
 |||||
 Db 335 IKTKRSI 341

RESULT 15
 TBA5_CHICK STANDARD; PRT; 448 AA.
 ID P09644;
 AC 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE TUBULIN ALPHA-5 CHAIN.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88296435; PubMed=3267229;
 RA Pratt L.F., Cleveland D.W.;
 RT "A survey of the alpha-tubulin gene family in chicken: unexpected
 RT sequence heterogeneity in the polypeptides encoded by five expressed
 RT genes.";
 RL EMBO J. 7:931-940(1988).
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
 CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
 CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
 CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
 CC -1- MISCELLANEOUS: THIS TUBULIN DOES NOT HAVE A C-TERMINAL TYROSINE.
 CC -1- MISCELLANEOUS: THERE ARE AT LEAST SEVEN ALPHA TUBULIN GENES
 CC (ALPHA-1 TO ALPHA-6, AND ALPHA-8), AND A PSEUDOGENE (ALPHA-7) IN
 CC CHICKEN.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X08061; CAA30852.1; ALT_INIT.
 DR PIR: S00470; UBCHA5.
 DR InterPro: IPR000217; Tubulin.
 DR InterPro: IPR003008; Tubulin_Ftsz.
 DR Pfam: PF00091; tubulin_1.
 DR PRINTS: PRO1161; TUBULIN.
 DR PRINTS: PRO1162; ALPHATUBULIN.
 DR PROSITE: PS00227; TUBULIN; 1.
 KM Microtubules; GTP-binding; Multigene family.

FT NP_BIND 142 148 GTP (POTENTIAL).
 SQ SEQUENCE 448 AA; 49948 MW; 2E498A2B20D7CBA6 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 448;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 114 IKTKRSI 120
 |||||
 Db 335 IKTKRSI 341

Search completed: May 9, 2002, 13:05:54
 Job time: 609 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:56:34 ; Search time 44.37 Seconds

(without alignments)
758.827 Million cell updates/sec

Title: US-09-787-083-8

Perfect score: 442
Sequence: 1 MKVSLSTLTSLTLPCEFLA.....YNHEATSPGVGLMDNMGL 442

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	2.0	382	2 E81195	phospholipase A1,
2	9	2.0	409	2 H81831	probable phospholip
3	8	1.8	286	2 B36971	outer membrane pho
4	8	1.8	299	2 T47989	RAV-like protein -
5	8	1.8	1396	2 S36851	L-shaped tail fibre
6	7	1.6	122	2 E75377	hypothetical prote
7	7	1.6	158	2 E86498	pts iia protein [i
8	7	1.6	158	2 E72124	tubulin alpha-1 ch
9	7	1.6	197	2 A54506	hypothetical prote
10	7	1.6	219	2 T19438	tubulin alpha chai
11	7	1.6	240	2 A61544	coat protein - pot
12	7	1.6	264	2 A44859	coat protein - pot
13	7	1.6	267	2 A60366	genome polyprotein
14	7	1.6	267	2 S14001	capsid protein - p
15	7	1.6	267	2 S26630	coat protein - pot
16	7	1.6	267	2 JCI527	tail fiber protein
17	7	1.6	267	2 S13339	abc transport ATP-
18	7	1.6	270	2 S73734	transaminase B hom
19	7	1.6	273	2 C73238	probable phosphotr
20	7	1.6	288	2 T27985	hypothetical prote
21	7	1.6	306	2 T27985	probable RAV-like
22	7	1.6	310	2 T02200	polyprotein - pota
23	7	1.6	313	2 JT0960	conserved hypochet
24	7	1.6	321	2 H82240	hypothetical prote
25	7	1.6	324	2 T23876	protease inhibitor
26	7	1.6	326	2 A43939	hypothetical prote
27	7	1.6	338	2 F69437	pectate lyase-like
28	7	1.6	341	2 T47653	hypothetical prote
29	7	1.6	342	2 G64411	hypothetical prote

30	7	1.6	342	2 C64394	hypothetical prote
31	7	1.6	349	2 T43920	yfc protein [impo
32	7	1.6	370	2 G83219	probable dihydroli
33	7	1.6	379	2 S13556	genome polyprotein
34	7	1.6	381	2 G82132	ribonuclease D VCI
35	7	1.6	382	2 S13237	tail fiber protein
36	7	1.6	384	2 S70638	tubulin alpha chai
37	7	1.6	384	2 S70641	tubulin alpha chai
38	7	1.6	384	2 S70640	tubulin alpha chai
39	7	1.6	387	2 S33517	tubulin alpha chai
40	7	1.6	407	2 T19895	hypothetical prote
41	7	1.6	415	2 H70477	UDP-N-acetylmuramo
42	7	1.6	419	2 F85064	hypothetical prote
43	7	1.6	419	2 T25577	hypothetical prote
44	7	1.6	423	1 UBRYA	tubulin alpha-1 ch
45	7	1.6	423	2 H70310	ammonium transport

ALIGNMENTS

RESULT 1

E81195 phospholipase A1, probable NMB0464 [imported] - Neisseria meningitidis (strain MC58 s

C:Species: Neisseria meningitidis

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C:Accession: E81195

R:Rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.

ri, H.; Qian, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Piatza, M.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappoli, R.

A>Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:2015755

A:Accession: E81195

A:Molecule type: DNA

A:Status: preliminary

A:Residues: 1382 <TET>

A:Cross-references: GB:AE002403; GB:AE002098; NID:g7225688; PIDN:AAF40901.1; PID:g722

A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NMB0464

Query Match 2.0%; Score 9; DB 2; Length 382;

Best Local Similarity 100.0%; Pred. No. 0.45;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIFLTQPV 287

Db 221 PEIFLTQPV 229

RESULT 2

H81831 probable phospholipase NMA2021 [imported] - Neisseria meningitidis (strain Z2491 serog

C:Species: Neisseria meningitidis

C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C:Accession: H81831

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo

Nature 404, 502-506, 2000

A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491

A:Reference number: A81775; MUID:20222556

A:Accession: H81831

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1409 <PAR>

A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85240.1; PID:g738

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMA2021

Query Match 2.0%; Score 9; DB 2; Length 409;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEFLTQPV 287
DB 248 PEFLTQPV 256

RESULT 3
B36971
Outer membrane phospholipase A (EC 3.1.1.-) precursor - *Klebsiella pneumoniae*
C:Species: *Klebsiella pneumoniae*
C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 18-Jun-1999
C:Accession: B36971; S40129
R:Brook, R.G.P.M.; Brinkman, E.; van Boxtel, R.; Bekkers, A.C.A.P.A.; Verheij, H.M.; Tomm
J. *Bacteriol.* 176, 861-870, 1994
A:Title: Molecular characterization of enterobacterial pldA genes encoding outer membran
A:Reference number: A36971; MUID:94131966
A:Accession: B36971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-286 <BRO>
A:Cross-references: EMBL:X76901; NID:g436880; PIDN:CAA54223.1; PID:g436881
A:Note: authors translated the codon AAG for residue 112 as Arg
C:Genetics:
A:Gene: pldA
C:Superfamily: bacterial phospholipase A1
C:Keywords: carboxylic ester hydrolase

Query Match 1.8%; Score 8; DB 2; Length 286;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 431 GVGGLMND 438
DB 277 GVGGLMND 284

RESULT 4
T47989
RAY-like protein - *Arabidopsis thaliana*
N:Alternate names: protein F21F14.140
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47989
R:Shotsune, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sa
submitted to the Protein Sequence Database, February 2000
A:Reference number: Z24481
A:Accession: T47989
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-299 <CHO>
A:Cross-references: EMBL:AL138642
A:Experimental source: cultivar Columbia; BAC clone F21F14
C:Genetics:
A:Map position: 3
A:Note: F21F14.140

Query Match 1.8%; Score 8; DB 2; Length 299;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 SLSFDDR 181
DB 292 SLSFDDR 299

RESULT 5
S36851

L-shaped tail fiber protein - phage T5
N:Alternate names: ltf protein
C:Species: phage T5
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 17-Nov-2000
C:Accession: S65934; S01984; S36851
R:Kallman, A.V.; Kulshin, V.E.; Shlyapnikov, M.G.; Ksenzenko, V.N.; Kryukov, V.M.
FEBS Lett. 366, 46-48, 1995
A:Title: The nucleotide sequence of the bacteriophage T5 ltf gene.
A:Reference number: S65934; MUID:95309401
A:Accession: S65934
A:Molecule type: DNA
A:Residues: 1-1396 <KAL>
A:Cross-references: EMBL:X69460; NID:g15415; PIDN:CAA49220.1; PID:g15416
R:Kallman, A.V.; Kryukov, V.M.; Bayev, A.A.
Nucleic Acids Res. 16, 6230, 1988
A:Title: The nucleotide sequence of bacteriophage T5 DNA at the region between early
A:Reference number: S01982; MUID:88289370
A:Accession: S01984
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 934-985, 'A', 987-1396 <KA2>
A:Cross-references: EMBL:X07559
C:Genetics:
A:Gene: ltf
C:Keywords: late protein; tail fiber

Query Match 1.8%; Score 8; DB 2; Length 1396;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 QSATQSAS 61
DB 87 QSATQSAS 94

RESULT 6
E75377
hypothetical protein - *Deinococcus radiodurans* (strain R1)
C:Species: *Deinococcus radiodurans*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: E75377
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
A:Reference number: A75250; MUID:20036896
A:Accession: E75377
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-122 <WHI>
A:Cross-references: GB:AE002002; GB:AE00513; NID:g6459345; PIDN:AAF11150.1; PID:g645
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRI580
A:Map position: 1

Query Match 1.6%; Score 7; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 EHEPELY 76
DB 115 EHEPELY 121

RESULT 7
E86498
Pta ITA Protein [imported] - *Chlamydomonas reinhardtii* (strain J138)
C:Species: *Chlamydomonas reinhardtii*
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001

C:Accession: E86498
 R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
 Nucleic Acids Res. 28, 2311-2314, 2000
 A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.
 A:Reference number: A86491; MUID:20330349
 A:Accession: E86498
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-158 <STO>
 A:Cross-references: GB:BA000008; NID:98978434; PIDN:BA98271.1; GSPDB:GN00142
 A:Experimental source: strain J138
 C:Genetics:
 A:Gene: ptnN.1
 C:superfamily: phosphotransferase system enzyme II; phosphotransferase system mannitol-5

Query Match 1.6%; Score 7; DB 2; Length 158;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 LSTLTLS 11
 DB 126 LSTLTLS 132

RESULT 8
 E72124
 pts Ila protein - Chlamydia pneumoniae (strain CWL029)
 C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
 C:Accession: E72124
 R:Rahman, S.; Mitchell, W.; Marathe, R.; Iammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
 Nature Genet. 21, 385-389, 1999
 A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
 A:Reference number: A72000; MUID:92206606
 A:Accession: E72124
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-158 <ARN>
 A:Cross-references: GB:AE001591; GB:AE001363; NID:94376311; PIDN:AAD18213.1; PID:9437631
 A:Experimental source: strain CWL029
 C:Genetics:
 A:Gene: ptnN.1
 C:superfamily: phosphotransferase system enzyme II; phosphotransferase system mannitol-5
 F:28-156/Domain: phosphotransferase system mannitol-specific enzyme II factor III homolog

Query Match 1.6%; Score 7; DB 2; Length 158;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 LSTLTLS 11
 DB 126 LSTLTLS 132

RESULT 9
 A54506
 tubulin alpha-1 chain - Plasmodium yoelii (fragment)
 C:Species: Plasmodium yoelii
 C:Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 13-Aug-1999
 C:Accession: A54506
 R:AKella, R.; Arasu, P.; Valdiva, A.B.
 Mol. Biochem. Parasitol. 30, 165-174, 1988
 A:Title: Molecular clones of alpha-tubulin genes of Plasmodium yoelii reveal an unusual
 A:Reference number: A54506; MUID:89014607
 A:Accession: A54506
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-197 <AKE>
 A:Cross-references: GB:M29816; NID:q160729; PIDN:AAA29779.1; PID:q160730
 C:superfamily: tubulin

Query Match 1.6%; Score 7; DB 2; Length 197;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120
 DB 82 IKTKRSI 88

RESULT 10
 T19438
 hypothetical protein C25A1.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T19438
 R:Mortimore, B.
 submitted to the EMBL Data Library, October 1996
 A:Reference number: Z19124
 A:Accession: T19438
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-219 <WIL>
 A:Cross-references: EMBL:Z81038; PIDN:CAB02762.1; GSPDB:GN00019; CESP:C25A1.1
 A:Experimental source: clone C25A1
 C:Genetics:
 A:Gene: CESP:C25A1.1
 A:Map position: 1
 A:introns: 33/1; 66/1; 151/2

Query Match 1.6%; Score 7; DB 2; Length 219;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 158 KDAKOLE 164
 DB 52 KDAKOLE 58

RESULT 11
 A61544
 tubulin alpha chain - common octopus (fragment)
 C:Species: Octopus vulgaris (common octopus)
 C:Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 10-Jul-1998
 C:Accession: A61544
 R:Zinov'eva, R.D.; Aleinikova, K.S.; Tomarev, S.I.
 Dokl. Akad. Nauk SSSR 302, 462-467, 1988
 A:Title: Isolation and structural characterization of cDNAs coding for alpha-tubulin
 A:Reference number: A61544
 A:Accession: A61544
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-240 <ZIN>
 C:superfamily: tubulin
 F:239-240/Cleavage site: Glu-Tyr (tubulin-specific carboxypeptidase) #status predicte
 F:239-240/Cross-link: peptide (Glu-Tyr) (by tubulin-tyrosine ligase) #status predicte

Query Match 1.6%; Score 7; DB 2; Length 240;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120
 DB 124 IKTKRSI 130

RESULT 12
 A44959
 coat protein - potato virus Y (strain N) (fragment)
 C:Species: potato virus Y, PVY
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Nov-2000

C:Accession: A44959
R:Hay, J.M.; Fellowes, A.P.; Timmerman, G.M.
Arch. Virol. 107, 111-122, 1989
A:Title: Nucleotide sequence of the coat protein gene of a necrotic strain of potato virus Y
A:Reference number: A44959; MUID:90025790
A:Accession: A44959
A:Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-264 <HAY>
A:Cross-references: GB:M22470; NID:9333508; PIDN:AAA47182.1; PID:9333509
C:Superfamily: tobacco etch virus genome polypeptide

Query Match 1.6%; Score 7; DB 2; Length 264;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKQ 162
|||||||
DB 8 TKKDAKQ 14

RESULT 13

A60366
coat protein - potato virus Y (strain TH)
C:Species: potato virus Y, PVY
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Nov-2000
C:Accession: A60366
R:Hattaya, T.; Sano, T.; Ohshima, K.; Shikata, E.
Virus Genes 4, 339-350, 1990
A:Title: Polymerase chain-reaction-mediated cloning and expression of the coat protein gene
A:Reference number: A60366; MUID:91143125
A:Accession: A60366
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-267 <HAY>
C:Superfamily: tobacco etch virus genome polypeptide

Query Match 1.6%; Score 7; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKQ 162
|||||||
DB 11 TKKDAKQ 17

RESULT 14

S14001
genome polypeptide - potato virus Y (strain H) (fragment)
N:Contains: coat protein
C:Species: potato virus Y, PVY
A:Variety: strain H
C:Date: 21-Nov-1993 #sequence_revision 08-Nov-1996 #text_change 17-Nov-2000
C:Accession: S14001
R:Dalmay, T.; Balazs, E.
Nucleic Acids Res. 18, 6721, 1990
A:Title: Nucleotide sequence of an altered virulence potato virus Y coat protein gene (F)
A:Reference number: S14001; MUID:91067494
A:Accession: S14001
A:Status: translation not shown
A:Molecule type: genomic RNA
A:Residues: 1-267 <DAL>
A:Cross-references: EMBL:X54611; NID:961449; PIDN:CAA38432.1; PID:9930257
A:Experimental source: strain H
C:Superfamily: tobacco etch virus genome polypeptide
C:Keywords: coat protein; polypeptide
F:1-267/Product: coat protein #status predicted <CP>

Query Match 1.6%; Score 7; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 42;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 156 TKKDAKQ 162
|||||||
DB 11 TKKDAKQ 17

RESULT 15

S26630
capsid protein - potato virus Y (fragment)
C:Species: potato virus Y, PVY
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 17-Nov-2000
C:Accession: S26630
R:Sudarsono, X.Y.Z.; Woloshuk, S.L.; Lommel, S.A.; Xiong, Z.; Hellman, G.M.; Wernsman
submitted to the EMBL Data Library, September 1992
A:Description: Nucleotide sequence of the capsid protein cistrons from six potato virus
A:Reference number: S26628
A:Accession: S26630
A:Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-267 <SUD>
A:Cross-references: EMBL:X68223; NID:961435; PIDN:CAA48303.1; PID:961436
C:Superfamily: tobacco etch virus genome polypeptide

Query Match 1.6%; Score 7; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKQ 162
|||||||
DB 11 TKKDAKQ 17

Search completed: May 9, 2002, 12:56:35
Job time: 605 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 13:05:52 ; Search time 28.06 Seconds

(Without alignments)
577.543 Million cell updates/sec

Title: US-09-787-083-6

Sequence: 442
1 MKVSLSTLTSLSCFAILA.....YNHEATSPGVGLMDMNGL 442

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 100059 seqs, 36664827 residues

Word size : 0

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	1.8	286	1	PAL_KLEPN
2	8	1.8	1396	1	PAL_KLEPN
3	7	1.6	186	1	PRL_ICTPU
4	7	1.6	197	1	TBA_PLAVO
5	7	1.6	237	1	TRMD_BUCAI
6	7	1.6	240	1	TBA_OCTYU
7	7	1.6	270	1	Y304_MYCPN
8	7	1.6	273	1	ILVE_THEMA
9	7	1.6	322	1	MIAA_PSEPU
10	7	1.6	326	1	SPI_BACBR
11	7	1.6	342	1	FIAD_METJA
12	7	1.6	342	1	Y755_METJA
13	7	1.6	347	1	NU2M_CAPII
14	7	1.6	347	1	NU2M_HUMAN
15	7	1.6	347	1	NU2M_PANTR
16	7	1.6	347	1	NU2M_PONPA
17	7	1.6	347	1	NU2M_SHEEP
18	7	1.6	423	1	AMT_AQUAE
19	7	1.6	443	1	TBA_ONCKE
20	7	1.6	448	1	TBA_HUMAN
21	7	1.6	448	1	TBA_HUMAN
22	7	1.6	449	1	TBA_CHICK
23	7	1.6	449	1	TBA_DROME
24	7	1.6	449	1	TBA_PHYPO
25	7	1.6	449	1	TBA_XENLA
26	7	1.6	450	1	TBAD_PHYPO
27	7	1.6	450	1	TBA_HARCO
28	7	1.6	450	1	TBA_NOTVI
29	7	1.6	451	1	TBA_CRIGR
30	7	1.6	451	1	TBA3_HOMAM
31	7	1.6	451	1	TBA_OCTDO
32	7	1.6	451	1	TBA_TORMA
33	7	1.6	511	1	LGT_CITUN

34	7	1.6	513	1	ATPA_HAEIN	P43714 haemophilus
35	7	1.6	536	1	60IM_BUCAP	P29431 buchnera ap
36	7	1.6	544	1	GP10_DICDI	O06885 dictyostell
37	7	1.6	567	1	GPV_RAT	O08770 rattus norv
38	7	1.6	610	1	FIMB_DICDI	P54680 dictyostell
39	7	1.6	768	1	YB23_HUMAN	O94117 homo sapien
40	7	1.6	783	1	FYB_HUMAN	O15117 homo sapien
41	7	1.6	900	1	AXNI_HUMAN	O15169 homo sapien
42	7	1.6	947	1	LKTA_PASSP	P55123 pasteurella
43	7	1.6	1026	1	VG37_BPT4	P03770 bacterioph
44	7	1.6	3061	1	POLG_PVYHU	O02963 p genome po
45	6	1.4	33	1	BR2B_RANES	P40838 rana escul

ALIGNMENTS

RESULT	1	STANDARD	PRT	286 AA.
PAL_KLEPN				
ID PAL_KLEPN				
AC P37446				
DT 01-OCT-1994 (Rel. 30, Last sequence update)				
DT 01-OCT-1994 (Rel. 30, Last sequence update)				
DT 01-NOV-1997 (Rel. 35, Last annotation update)				
DE PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DEFERENT-RESISTANT				
DE PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-				
DE ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).				
GN PLDA.				
OS Klebsiella pneumoniae.				
OC Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;				
OC Klebsiella.				
OX NCBI_TaxID=573;				
RN [1]				
RP SEQUENCE FROM N.A.				
RX MEDLINE=94131966; PubMed=8300539;				
RA Brok R.G.P.M., Brinkman E., van Bortel R., Bekkers A.C.A.P.,				
RA Verheij H.M., Tomassen J.;				
RT Molecular characterization of enterobacterial plda genes encoding				
RT outer membrane phospholipase A.;				
RL J. Bacteriol. 176:861-870(1994).				
CC -1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE				
CC A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.				
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACLYGLYCERO-				
CC PHOSPHOCHOLINE + A FATTY ACID ANION.				
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACLYGLYCERO-				
CC PHOSPHOCHOLINE + A FATTY ACID ANION.				
CC -1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.				
CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES				
CC LOCATED THERE.				
CC -----				
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration				
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC the European Bioinformatics Institute. There are no restrictions on its				
CC use by non-profit institutions as long as its content is in no way				
CC modified and this statement is not removed. Usage by and for commercial				
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC or send an email to license@isb-sib.ch).				
CC -----				
CC EMBL: X76901; CAA54223.1; -				
DR PIR: B36971; B36971.				
DR PIR: S40129; S40129.				
DR InterPro: IPR003187; PLA1.				
DR Pfam: PF02253; PLA1; 1.				
KW Hydrolase; Lipid degradation; Outer membrane; signal; Calcium.				
FT SIGNL 1 20				
FT CHAIN 21 286				
FT ACT_SITE 161 161				
FT ACT_SITE 286 AA; 32544 MM; 3E39F863085108A3 CRC64;				
SQ SEQUENCE				

Query Match 1.8%; Score 8; DB 1; Length 286;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 431 GVGGLMIND 438
 DB 277 GVGGLMIND 284

RESULT 2

VLTFF_BPRT5 STANDARD; PRT: 1396 AA.
 AC P13390; 048502;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE L-SHAPED TAIL FIBER PROTEIN (LTF PROTEIN).
 GN LTF.
 OS Bacteriophage T5.
 OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae.
 OX NCBI_TaxID=10726;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95309401; PubMed=7789514;
 RA Kallman A.V., Kulshin V.E., Shlyapnikov M.G., Ksenzenko V.N.,
 RA Krynukov V.M.;
 RT "The nucleotide sequence of the bacteriophage T5 ltf gene.";
 RT FEBS Lett. 366:46-48(1995).
 RL [2]
 RN SEQUENCE FROM N.A.
 RA Kallman A.V.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
 RX MEDLINE=88289370; PubMed=3267228;
 RA Kallman A.V., Krynukov V.M., Bayev A.A.;
 RT "The nucleotide sequence of bacteriophage T5 DNA at the region
 RT between early and late genes.";
 RT Nucleic Acids Res. 16:6230-6230(1988).
 RL [1]
 CC -1- FUNCTION: NONESSENTIAL PROTEIN THAT MEDIATES BINDING TO THE
 CC POLYMANNOSE O ANTIGEN.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X69460; CAA49220.1; -
 DR EMBL: AJ001191; CAA04591.1; -
 DR PIR: S01982; S01982.
 KW Late protein.
 FT CONFICIT 986 986 V -> A (IN REF. 2).
 SQ SEQUENCE 1396 AA; 147989 MW; 18CD2192F65FFC1 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 1396;
 Best Local Similarity 100.0%; Pred. No. 7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 54 QSATOSAS 61
 DB 87 QSATOSAS 94

RESULT 3

PRL_ICTPU STANDARD; PRT: 186 AA.
 AC P51904;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE PROLACTIN (PRL).
 OS Ictalurus punctatus (Channel catfish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Siluriformes; Ictaluridae; Ictalurus.
 OX NCBI_TaxID=7998;
 RN [1]
 RP SEQUENCE.

RC TISSUE-Pituitary;
 RX MEDLINE=93364578; PubMed=1308206;
 RA Watanabe K., Igarashi A., Noso T., Chen T.T., Dunham R.A.,
 RA Kawachi H.;
 RT "Chemical identification of catfish growth hormone and prolactin.";
 RT MOL. Mar. Biol. Biotechnol. 1:239-249(1992).
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
 DR HSP: Q28632; IAN3.
 DR InterPro: IPR001400; SOMATOTROPIN.
 DR Pfam: PF00103; hormone; 1.
 DR PRINTS: PR00836; SOMATOTROPIN.
 DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
 DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
 KW Hormone; Pituitary.
 FT DISULFID 45 159 BY SIMILARITY.
 FT DISULFID 176 186 BY SIMILARITY.
 SQ SEQUENCE 186 AA; 20606 MW; 7BCCDB9718F44E74 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 186;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 41 NDIGDON 47
 DB 140 NDIGDON 146

RESULT 4

TBA_PLAYO STANDARD; PRT: 197 AA.
 AC P12543;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE TUBULIN ALPHA CHAIN (FRAGMENT).
 OS Plasmodium berghei yoelli.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5862;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89014607; PubMed=2459618;
 RA Akella R., Arasu P., Valodia A.B.;
 RT "Molecular clones of alpha-tubulin genes of Plasmodium yoelli reveal
 RT an unusual feature of the carboxy terminus.";
 RT Mol. Biochem. Parasitol. 30:165-174(1988).
 RL [1]
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
 CC BINDS TWO MOLES OF GTP. ONE AT AN EXCHANGEABLE SITE ON THE BETA
 CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
 CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

EMBL: M29816; AA29779.1; -
 DR PIR: A54506; A54506.
 DR InterPro: IPR000217; Tubulin.
 DR InterPro: IPR003008; Tubulin_ftsz.
 DR Pfam: PF00091; tubulin; 1.
 DR PROSITE: PS00227; TUBULIN; PARTIAL.

KW Microtubules; GMP-binding.
 FT NON_TER 1
 SQ SEQUENCE 197 AA; 22111 MW; 4DAAF199CCGCCD319 CRC64;

Query Match
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120
 |||||
 Db 82 IKTKRSI 88

RESULT 5
 TRMD_BUCAL STANDARD; PRT; 237 AA.

AC P57476;
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE tRNA (GUANINE-N1)-METHYLTRANSFERASE (EC 2.1.1.31) (MIG-
 DE METHYLTRANSFERASE) (tRNA [GM37] METHYLTRANSFERASE).
 GN TRMD OR BU396.
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
 symbiotic bacterium).
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 OX NCBI_TaxID=118099;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TOKYO 1998;
 RX MEDLINE=20445173; PubMed=10993077;
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
 RT "Genome sequence of the endocellular bacterial symbiont of aphids
 Buchnera sp. APS.";
 RL Nature 407:81-86(2000).
 CC -1- FUNCTION: SPECIFICALLY METHYLATES GUANOSINE-37 IN VARIOUS TRNAS
 CC (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + TRNA = S-ADENOSYL-
 CC L-HOMOCYSTEINE + TRNA CONTAINING N(1)-METHYLGUANINE.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE RNA METHYLTRANSFERASE TRMD FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC -----
 CC EMBL: AP001119; BAB13099.1; -
 DR InterPro: IPR002649; tRNA_mig_MT.
 DR Pfam: PF01746; tRNA_mig_MT; 1.
 DR Prodom: PD004978; tRNA_mig_MT; 1.
 KW Transferase; Methyltransferase; tRNA processing; Complete proteome.
 SQ SEQUENCE 237 AA; 27335 MW; 86CE3836B35EC4AB CRC64;

Query Match
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LAIOQAK 25
 |||||
 Db 78 LAIOQAK 84

RESULT 6
 TBA_OCTVU STANDARD; PRT; 240 AA.
 AC P24635;
 DT 01-MAR-1992 (Rel. 21, Created)

DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE TUBULIN ALPHA CHAIN (FRAGMENT).
 OS Octopus vulgaris (Octopus).
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;
 OC Incirata; Octopodidae; Octopus.
 OX NCBI_TaxID=6645;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue-Lens;
 RA Zinov'Eva R.D., Aleinikova K.S., Tomarev S.I.;
 RT "Isolation and structural characterization of cDNAs coding for alpha-
 RT tubulin of the octopus eye lens.";
 RL Dokl. Akad. Nauk SSSR 302:462-467(1988).
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
 CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
 CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
 CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC -----
 CC EMBL: X15845; CAA33844.1; -
 DR PIR: A61544; A61544.
 DR InterPro: IPR000217; Tubulin.
 DR InterPro: IPR003008; Tubulin_FtsZ.
 DR Pfam: PF00091; tubulin; 1.
 DR PROSITE: PS00227; TUBULIN; PARTIAL.
 KW Microtubules; GMP-binding.
 FT NON_TER 1
 FT SITE 1
 FT SITE 240
 FT SITE 26961
 SQ SEQUENCE 240 AA; 26961 MW; 13BB3A1F740F2416 CRC64;

Query Match
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120
 |||||
 Db 124 IKTKRSI 130

RESULT 7
 Y304_MYCPN STANDARD; PRT; 270 AA.

AC P75355;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHEICAL ABC TRANSPORTER ATP-BINDING PROTEIN MG304 HOMOLOG
 DE (A05_ORF2701).
 GN MPN433 OR MP408.
 GN MPN433 OR MP408.
 OS Mycoplasma pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molluscites;
 OC Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RX MEDLINE=97105885; PubMed=8948633;
 RA Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
 RT pneumoniae.";
 RL Nucleic Acids Res. 24:4420-4449(1996).
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY

(ABC TRANSPORTERS).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AE000040; AAB96056.1; -
DR InterPro: IPR003593; AAA.
DR InterPro: IPR003439; ABC_transport.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00005; ABC_tran; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
DR Hypothetical protein: ATP-binding; Transport; Complete proteome.
FT NP_BIND 36 43 ATP (POTENTIAL).
SQ SEQUENCE 270 AA; 30770 MW; 6512640E4BC051B4 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 270;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 168 KQFTPLS 174
Db 125 KQFTPLS 131

RESULT 8
MIAA_PSEPU
ID ILVE_THEME STANDARD; PRT; 273 AA.
AC P74921;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROBABLE BRANCHED-CHAIN AMINO ACID AMINOTRANSFERASE (EC 2.6.1.42)
DE (BCAA).
DE ILVE OR TM0831.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxId=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; Pubmed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima."
RL Nature 399:323-329(1999).
RN [2]
RP SEQUENCE OF 68-273 FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=97017137; Pubmed=8863738;
RA Guldpaud O., Lapedan B., Forcette P.;
RT "A gyrB-like gene from the hyperthermophilic bacterium Thermotoga
RT maritima."
RL Gene 174:121-128(1996).
CC -1- CATALYTIC ACTIVITY: L-LEUCINE + 2-OXOGLOUTARATE -> 4-METHYL-2-
CC OXOBUTANOATE + L-GLUTAMATE (ALSO ACTS ON L-ISOLEUCINE AND
CC L-VALINE).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: VALINE AND ISOLEUCINE BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO CLASS-IV OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AE001750; AAD35913.1; -
DR EMBL: U49692; AAC44497.1; -
DR HSSP: P00510; IACG.
DR TIGR: TM0831; -
DR InterPro: IPR001544; Aminotran.4.
DR Pfam: PF01063; aminotran.4; 1.
DR ProDom: PD001961; Aminotran.4; 1.
DR PROSITE: PS00770; AA_TRANSFER_CLASS.4; 1.
DR Transferase: Aminotransferase; Branched-chain amino acid biosynthesis;
KW Pyridoxal phosphate; Complete proteome.
FT CONFLICT 77 77 A -> R (IN REF. 2).
SQ SEQUENCE 273 AA; 31158 MW; 2162B705612E90E3 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 273;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 172 PLSTLSPD 178
Db 61 PLSTLSPD 67

RESULT 9
MIAA_PSEPU
ID MIAA_PSEPU STANDARD; PRT; 322 AA.
AC Q30762;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE tRNA DELTA(2)-ISOPENTENYL-PYROPHOSPHATE TRANSFERASE (EC 2.5.1.8) (IPP
DE TRANSFERASE) (ISOPENTENYL-DIPHOSPHATE:TRNA ISOPENTENYLTRANSFERASE)
DE (IPPTASE) (IPPT).
OS MIAA.
GN Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxId=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M.
RA Olekhnovich I.N., Gussin G.N.;
RT "Attenuation of the pseudomonas putida tpe and ttpdc genes."
RL Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE FIRST STEP IN THE BIOSYNTHESIS OF
CC 2-METHYLTHIO-N6-(DELTA(2)-ISOPENTENYL)-ADENOSINE (MSI[2]I[6]A)
CC ADJACENT TO THE ANTICODON OF SEVERAL TRNA SPECIES (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ISOPENTENYL DIPHOSPHATE + TRNA -
CC PYROPHOSPHATE + TRNA CONTAINING 6-ISOPENTENYLADENOSINE.
CC -1- SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AF016312; AAB69443.1; -
DR InterPro: IPR002627; IPPT.
DR Pfam: PF01715; IPPT; 1.
DR ProDom: PD004674; IPPT; 1.
DR Transferase: Nucleotidyltransferase; tRNA processing; ATP-binding.
FT NP_BIND 12 19 ATP (POTENTIAL).

```

SQ SEQUENCE 322 AA: 35488 MW: DIA81A2D7B32F6A CRC64:
Query Match 1.6%; Score 7: DB 1; Length 322;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 FVDEVRS 39
    |||||
Db 231 FVDEVRS 237

RESULT 10
SPL_BACBR STANDARD; PRT: 326 AA.
ID SPL_BACBR
AC P4313;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PROTEASE INHIBITOR PRECURSOR (BBRP1).
OS Bacillus brevis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Brevibacillus.
ON NCBI_TaxID=1393;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-32; 104-112 AND 122-136.
RC STRAIN=HPD31;
RX MEDLINE=92304060; PubMed=1610177;
RA Shiga Y., Hasegawa K., Tsuboi A., Yamagata H., Ueda S.;
RT "Characterization of an extracellular protease inhibitor of Bacillus
RT brevis HPD31 and nucleotide sequence of the corresponding gene.";
RL Appl. Environ. Microbiol. 58:525-531(1992).
CC -1- FUNCTION: SHOWS INHIBITORY ACTIVITY TOWARDS SERINE PROTEASES, SUCH
CC AS TRYPSIN, CHYMOTRYPSIN, AND SUBTILISIN. MAY FORM A TRYPSIN-
CC INHIBITOR COMPLEX IN A MOLAR RATIO OF 1:1. IT IS HEAT RESISTANT AT
CC NEUTRAL AND ACIDIC PHs.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- PTM: MAY UNDERGO ACTIVATION AFTER SECRETION. IT IS PRODUCED.
CC EXTRACELLULARLY IN MULTIPLE FORMS HAVING AT LEAST THREE DIFFERENT
CC MOLECULAR WEIGHTS (BBRP1-A, -B, AND -C).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D10696; BAA01538.1; -
KW Serine protease inhibitor; Protease inhibitor; Signal; Repeat.
FT SIGNAL 1 24
FT CHAIN 25 326 SERINE PROTEASE INHIBITOR.
FT CHAIN 104 326 SERINE PROTEASE INHIBITOR-C.
FT CHAIN 104 ? SERINE PROTEASE INHIBITOR-B (POTENTIAL).
FT CHAIN 122 326 SERINE PROTEASE INHIBITOR-A.
FT DOMAIN 177 304 CONTAINS TWO APPROXIMATE REPEATS.
FT REPEAT 177 208 1.
FT REPEAT 272 304 2.
SO SEQUENCE 326 AA; 35100 MW; 1C0456ABFA912F77 CRC64;

Query Match 1.6%; Score 7: DB 1; Length 326;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 PVAFVDE 36
    |||||
Db 119 PVAFVDE 125

RESULT 11
FLAD_METJA

SQ SEQUENCE 342 AA: 39950 MW: B384DDE1775566C CRC64:
Query Match 1.6%; Score 7: DB 1; Length 342;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 356 PDIDY 362
    |||||
Db 255 PDIDY 261

RESULT 12
Y755_METJA STANDARD; PRT: 342 AA.
ID Y755_METJA
AC O58165;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN M0755 PRECURSOR.
GN M0755.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
ON NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Usterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Clifton M.D., Roberts K.M., Hurst M.A., Kalne B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
CC -1- SIMILARITY: STRONG. TO M.VOLTAE FLAD, ALSO TO FLAE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67533; AAB98898.1; -
DR TIGR: M08093; -
KW Flagella; Complete proteome.
SO SEQUENCE 342 AA; 39950 MW; B384DDE1775566C CRC64;

Query Match 1.6%; Score 7: DB 1; Length 342;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 356 PDIDY 362
    |||||
Db 255 PDIDY 261

```

RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Colton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Kent H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, *Methanococcus*
 RT *jannaschii*."; Science 273:1058-1073(1996).

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U67521; AAB98748.1; -
 DR TIGR: M10755; -
 KW Hypothetical protein; Signal; Complete proteome.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 342 HYPOTHETICAL PROTEIN M10755.
 SQ SEQUENCE 342 AA; 38778 MW; BD25220A7EAD85C CRC64;

Query Match 1.6%; Score 7; DB 1; Length 342;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 76 YTTALEN 82
 |||||
 Db 90 YTTALEN 96

RESULT 13
 NU2M_CAPHI STANDARD; PRT; 347 AA.
 ID NU2M_CAPHI
 AC 036346;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 2 (EC 1.6.5.3).
 GN MTND2 OR ND2.
 OS Capra hircus (Goat).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Capra.
 OX NCBI_TaxID=9925;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA DOVC P., Mann W., Hecht W.;
 RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE -> NAD(+) + UBIQUINOL.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 CC INNER MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 2 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X72965; CAAS1468.1; -
 DR InterPro: IPR003917; NADHdb_oxdrdctse2.
 DR InterPro: IPR001750; Oxidored_q1.
 DR Pfam: PF00361; oxidored_q1; 1.
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
 SQ SEQUENCE 347 AA; 39136 MW; 6129CFF90980475E CRC64;

Query Match 1.6%; Score 7; DB 1; Length 347;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 8 LTLSILS 14
 |||||
 Db 155 LTLSILS 161

RESULT 14
 NU2M_HUMAN STANDARD; PRT; 347 AA.
 ID NU2M_HUMAN
 AC P03891; Q34769; Q9TGI1; Q9TGI2; Q9TGI3; Q9TGI4;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 2 (EC 1.6.5.3).
 GN MTND2 OR ND2.
 OS Homo sapiens (Human).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81173052; PubMed=7219534;
 RA Anderson S., Bankier A.T., Barrell B.G., de Bruijn M.H.L.,
 RA Coulson A.R., Drouin J., Eperon I.C., Nierlich D.P., Roe B.A.,
 RA Sanger F., Schreier P.H., Smith A.J.H., Staden R., Young I.G.;
 RT "Sequence and organization of the human mitochondrial genome."; Nature 290:457-465(1981).
 RL Nature 290:457-465(1981).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81170577; PubMed=6260957;
 RA Sanger F., Coulson A.R., Barrell B.G., Smith A.J.H., Roe B.A.;
 RT "Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing."; J. Mol. Biol. 143:161-178(1980).
 RL J. Mol. Biol. 143:161-178(1980).
 RN [3]
 RP SEQUENCE FROM N.A., AND VARIANTS.
 RX MEDLINE=9813898; PubMed=9475751;
 RA Wise C.A., Strahl M., Eastaugh S.;
 RT "Departure from neutrality at the mitochondrial NADH dehydrogenase subunit 2 gene in humans, but not in chimpanzees."; Genetics 148:409-421(1998).
 RL Genetics 148:409-421(1998).
 RN [4]
 RP SEQUENCE FROM N.A., AND VARIANTS ILE-43; LEU-325 AND THR-331.
 RC TISSUE=Placenta;
 RX MEDLINE=95132634; PubMed=7530363;
 RA Horai S., Hayasaka K., Kondo R., Tsugane K., Takahata N.;
 RT "Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial DNAs."; Proc. Natl. Acad. Sci. U.S.A. 92:532-536(1995).
 RL Proc. Natl. Acad. Sci. U.S.A. 92:532-536(1995).
 RN [5]
 RP IDENTIFICATION OF PROTEIN.
 RX MEDLINE=85188293; PubMed=3921850;
 RA Chomyn A., Mariotti P., Cleeter M.W.J., Ragan C.I., Matsuno-Yagi A.,
 RA Hattel Y., Doollittle R.F., Attardi G.;
 RT "Six unidentified reading frames of human mitochondrial DNA encode components of the respiratory-chain NADH dehydrogenase."; Nature 314:592-597(1985).
 RL Nature 314:592-597(1985).
 RN [6]
 RP VARIANT LHON ASP-150.
 RX MEDLINE=91144615; PubMed=1900003;
 RA Johns D.R., Berman J.;
 RT "Alternative, simultaneous complex I mitochondrial DNA mutations in Leber's hereditary optic neuropathy."; Biochem. Biophys. Res. Commun. 174:1324-1330(1991).
 RL [7]
 RP VARIANTS L-42; R-63; A-119; P-148; S-150; T-159 AND A-185.
 RX MEDLINE=92098084; PubMed=1757091;
 RA Marzuki S., Noer A.S., Lertit P., Thyagarajan D., Kapsa R.,
 RA Utthanaphol P., Byrne E.;

RT "Normal variants of human mitochondrial DNA and translation products:
RT the building of a reference data base."; Hum. Genet. 88:139-145(1991).
RN [9]
RX MEDLINE-92120513; PubMed-1732158;
RP VARIANT LHON SER-259.
RA Brown M.D., Voljavec A.S., Lott M.T., Torroni A., Yang C.C.,
RA Wallace D.C.;
RT "Mitochondrial DNA complex I and III mutations associated with
RT Leber's hereditary optic neuropathy."; Genetics 130:163-173(1992).
RN [9]
RP VARIANT AD SER-331.
RX MEDLINE-92118019; PubMed-1370613;
RA Lin F.-H., Lin R., Wisniewski H.M., Hwang Y.-W., Grundke-Iqbal I.,
RA Healy-Louie G., Iqbal K.;
RT "Detection of point mutations in codon 331 of mitochondrial NADH
RT dehydrogenase subunit 2 in Alzheimer's brains."; Biochem. Biophys. Res. Commun. 182:238-246(1992).
RN [10]
RP VARIANT THR-57.
RX MEDLINE-98127994; PubMed-9461455;
RA Rieder M.J., Taylor S.L., Tobe V.O., Nickerson D.A.;
RT "Automating the identification of DNA variations using quality-based
RT fluorescence re-sequencing: analysis of the human mitochondrial
RT genome."; Nucleic Acids Res. 26:967-973(1998).
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE.
CC -1- DISEASE: DEFECTS IN MTND2 ARE ONE OF THE CAUSES OF LEBER'S
CC HEREDITARY OPTIC NEUROPATHY (LHON). A MATERNALLY INHERITED DISEASE
CC RESULTING IN ACUTE BILATERAL BLINDNESS DUE TO RETINAL DEGENERATION
CC PREDOMINANTLY IN YOUNG MEN. CARDIAC CONDUCTION DEFECTS AND
CC NEUROLOGICAL DEFECTS HAVE ALSO BEEN DESCRIBED. RESULTING IN OPTIC
CC NERVE DEGENERATION AND CARDIAC DYSRHYTHMIA.
CC -1- DISEASE: DEFECTS IN MTND2 COULD BE ASSOCIATED WITH ALZHEIMER'S
CC DISEASE (AD).
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: J01415; AAB58944.1; -;
DR EMBL: V00662; CAA24027.1; -;
DR EMBL: M10546; AAA65502.1; ALT_INIT.
DR EMBL: D38112; BAA07291.1; -;
DR EMBL: AF014882; AAC25441.1; -;
DR EMBL: AF014884; AAC25443.1; -;
DR EMBL: AF014885; AAC25444.1; -;
DR EMBL: AF014887; AAC25446.1; -;
DR EMBL: AF014889; AAC25448.1; -;
DR EMBL: AF014891; AAC25450.1; -;
DR EMBL: AF014892; AAC25451.1; -;
DR EMBL: AF014895; AAC25454.1; -;
DR EMBL: AF014896; AAC25455.2; -;
DR EMBL: AF014897; AAC25456.1; -;
DR EMBL: AF014898; AAC25457.1; -;
DR EMBL: AF014899; AAC25458.2; -;
DR EMBL: AF014900; AAC25459.1; -;
DR EMBL: AF014901; AAC25460.1; -;
DR PIR: A00414; DNHUN2.
DR MIM: 502500; -;
DR MIM: 516001; -;
DR MIM: 535000; -;
DR InterPro: IPR003917; NADHub_oxdrcdctse2.
DR InterPro: IPR001750; Oxidored_g1.

DR Pfam: PF00361; oxidored_g1.1.
KW Oxidoreductase; NAD: Ubiquinone; Mitochondrion; Transmembrane;
KW Disease mutation; Leber's hereditary optic neuropathy;
KW Alzheimer's disease; Polymorphism.
FT VARIANT 42 42
FT P -> L.
FT VARIANT 43 43
FT V -> I.
FT /FTid=VAR_008590.
FT /FTid=VAR_011348.
FT I -> T.
FT /FTid=VAR_008591.
FT Q -> R.
FT /FTid=VAR_008592.
FT I -> V.
FT /FTid=VAR_011349.
FT N -> S.
FT /FTid=VAR_011350.
FT T -> A.
FT /FTid=VAR_008593.
FT S -> P.
FT /FTid=VAR_008594.
FT N -> D (IN LHON; SECONDARY MUTATION; DOES
FT NOT SEEM TO DIRECTLY CAUSE THE DISEASE).
FT /FTid=VAR_004755.
FT N -> S.
FT /FTid=VAR_008595.
FT I -> T.
FT /FTid=VAR_008596.
FT T -> A.
FT /FTid=VAR_008597.
FT L -> M.
FT /FTid=VAR_011351.
FT G -> S (IN LHON; RARE PRIMARY MUTATION).
FT /FTid=VAR_004756.
FT A -> T.
FT /FTid=VAR_011352.
FT A -> V.
FT /FTid=VAR_011353.
FT I -> T.
FT /FTid=VAR_011354.
FT F -> L.
FT /FTid=VAR_011355.
FT A -> S (IN AD).
FT /FTid=VAR_004758.
FT A -> T.
FT /FTid=VAR_004757.
FT T -> A.
FT /FTid=VAR_011356.
FT SEQUENCE 347 AA; 38961 MW; C06FD982317C3F2D CRC64;
FT SQ
Query Match 1.6%; Score 7; DB 1; Length 347;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 8 LTLSTLS 14
Db 155 LTLSTLS 161
RESULT 15
NU2M_PANTR
ID NU2M_PANTR STANDARD; PRT; 347 AA.
AC 021798;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 2 (EC 1.6.5.3).
GN MTND2 OR ND2.
OS Pan troglodytes (Chimpanzee).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.
OX NCBI_Taxid=9598;

```
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:98133898; PubMed:9475751;
RA Wise C.A., Strahl M., Eastrel S.;
RT "Departure from neutrality at the mitochondrial NADH dehydrogenase
RL subunit 2 gene in humans, but not in chimpanzees.";
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF014908; AAC25467.1; -
DR EMBL: AF014909; AAC25468.1; -
DR EMBL: AF014910; AAC25469.1; -
DR EMBL: AF014911; AAC25470.1; -
DR EMBL: AF014912; AAC25471.1; -
DR EMBL: AF014913; AAC25472.1; -
DR EMBL: AF014914; AAC25473.1; -
DR EMBL: AF014915; AAC25474.1; -
DR EMBL: AF014916; AAC25475.1; -
DR EMBL: AF014917; AAC25476.1; -
DR EMBL: AF014918; AAC25477.1; -
DR EMBL: AF014919; AAC25478.1; -
DR EMBL: AF014920; AAC25479.1; -
DR EMBL: AF014921; AAC25480.1; -
DR InterPro: IPR003917; NADHUb_oxidrcdscse2.
DR InterPro: IPR001750; Oxidored_q1.
DR Pfam: PF00361; Oxidored_q1; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
FT VARIANT 94 94 P -> S (IN STRAIN A-292).
SQ SEQUENCE 347 AA; 39020 MW; 2E8269D105810D4E CRC64;
```

```
Query Match 1.6%; Score 7; DB 1; Length 347;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 8 LTLSTLS 14
| | | | | | |
DB 155 LTLSTLS 161
```

Search completed: May 9, 2002, 13:05:53
Job time: 608 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 13:05:18 : Search time 77.39 Seconds
(without alignments)
835,410 Million cell updates/sec

Title: US-09-787-083-6

Sequence: 442
1 MKVSLSTLTSLSCPAIA.....YNHEATSGVGLMDNMGL 442

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 473505 seqs, 146272329 residues

Word size : 0

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

- 1: SPREMBL_17:*
- 2: sp_Archaea:*
- 3: sp_Bacteria:*
- 4: sp_Fungi:*
- 5: sp_human:*
- 6: sp_invertebrate:*
- 7: sp_mammal:*
- 8: sp_mhbc:*
- 9: sp_mhbc:*
- 10: sp_mhbc:*
- 11: sp_plant:*
- 12: sp_rodent:*
- 13: sp_virus:*
- 14: sp_vertebrate:*
- 15: sp_vertebrate:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	2.0	382	2 09K007	09K007 neisseria m
2	9	2.0	409	2 09J721	09J721 neisseria m
3	9	2.0	862	2 09CL62	09CL62 pasteurella
4	9	2.0	1432	10 09AX85	09AX85 oryza sativ
5	8	1.8	299	10 09M268	09M268 arabidopsi
6	8	1.8	306	2 09CL22	09CL22 pasteurella
7	7	1.6	55	12 067608	067608 tomato goi
8	7	1.6	72	12 0912S5	0912S5 potato viru
9	7	1.6	72	12 0912S4	0912S4 potato viru
10	7	1.6	95	12 0912S3	0912S3 potato viru
11	7	1.6	96	12 0912S8	0912S8 potato viru
12	7	1.6	100	3 012148	012148 saccharomy
13	7	1.6	101	12 0912S6	0912S6 potato viru
14	7	1.6	116	2 09C1E9	09C1E9 lactococcus
15	7	1.6	122	2 09R016	09R016 delinococcus
16	7	1.6	124	12 0912S7	0912S7 potato viru
17	7	1.6	131	6 09GMD4	09GMD4 macaca fusc
18	7	1.6	138	12 0912O7	0912O7 potato viru
19	7	1.6	147	5 036035	036035 hexamita in

20	7	1.6	158	2 0929C1	0929C1 chlamydia p
21	7	1.6	169	2 09L2A7	09L2A7 streptomyces
22	7	1.6	189	13 09DFT2	09DFT2 notochemia
23	7	1.6	198	4 09BUX9	09BUX9 homo sapien
24	7	1.6	212	13 091819	091819 ictalurus p
25	7	1.6	214	5 09VTR5	09VTR5 drosophila
26	7	1.6	216	2 09X6H6	09X6H6 streptococ
27	7	1.6	218	2 09RF14	09RF14 streptococ
28	7	1.6	219	5 017594	017594 caenorhabdl
29	7	1.6	220	2 09XCK1	09XCK1 streptococ
30	7	1.6	220	5 036036	036036 spirochaet
31	7	1.6	221	12 065002	065002 alfalfa mos
32	7	1.6	227	5 09BWF4	09BWF4 glossina mo
33	7	1.6	233	6 09WZB0	09WZB0 ovis aries
34	7	1.6	236	13 09DPS8	09DPS8 notochemia
35	7	1.6	264	12 085260	085260 potato viru
36	7	1.6	265	9 038155	038155 bacterioph
37	7	1.6	267	9 099363	099363 bacterioph
38	7	1.6	267	12 085261	085261 potato viru
39	7	1.6	267	12 098628	098628 potato viru
40	7	1.6	267	12 086842	086842 potato viru
41	7	1.6	267	12 099026	099026 potato viru
42	7	1.6	267	12 005476	005476 potato viru
43	7	1.6	267	12 009711	009711 potato viru
44	7	1.6	267	12 09WH17	09WH17 potato viru
45	7	1.6	267	12 09QNM9	09QNM9 potato viru

ALIGNMENTS

RESULT 1
ID 09K007 PRELIMINARY: PRT: 382 AA.
AC 09K007;
DT 01-OCT-2000 (TREMREL. 15, Created)
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)
DT 01-JUN-2001 (TREMREL. 17, Last annotation update)
DE PHOSPHOLIPASE A1, PUTATIVE.
GN NMB0464.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
Eisen J.A., Ketchum K.A., Hood B.W., Peden J.F., Dodson R.J.,
Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey E.K.,
Hait D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
Mason T., Clecko A., Parksey D.S., Blair E., Ciftone H., Clark E.B.,
Cotton M.D., Uterback T.R., Khouri H., Qin H., Yamathavan J.,
Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,
Smith H.O., Fraser C.M., Moxon E.R., Rappelli R., Venter J.C.;
RA "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58.";
RL Science 287:1809-1815(2000).
DR EMBL: AE002403; AAP40901.1; -;
DR TIGR: NMB0464; -;
DR InterPro: IPR003187; PLAI.
DR Pfam: PF02253; PLAI; 1.
KW Complete proteome.
SQ SEQUENCE 382 AA; 42714 MW; B468A802F062E836 CRC64;

Query Match 2.0%; Score 9; DB 2; Length 382;
Best Local Similarly 100.0%; Pred.No. 1.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 279 PEIRFOPV 287
DB 221 PEIRFOPV 229

```

RESULT 2
O9JUT21 ID 09JUT21 PRELIMINARY: PRT: 409 AA.
AC 09JUT21:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PUTATIVE PHOPHOLIPASE.
GN MMA2021.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-22491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltham T., Hamlin N., Holtroyd S.,
RA Jagers K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis 22491."
RL Nature 404:502-506 (2000).
DR EMBL: AL162757; CAB85240.1;
DR InterPro: IPR003187; Pfam:
DR Pfam: PF02253; Pfam: 1.
KW Complete proteome.
SQ SEQUENCE 409 AA; 45862 MW; CD6585B064D01A41 CRC64;

Query Match
Best Local Similarity 2.0%; Score 9; DB 2; Length 409;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIFLTQPV 287
Db 248 PEIFLTQPV 256

RESULT 3
O9CL62 ID 09CL62 PRELIMINARY: PRT: 862 AA.
AC 09CL62:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CVGSI.
GN CVGSI OR PM1380.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang O., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida PM70."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465 (2001).
CC -1- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC REGULATORY COMPONENTS OF SENSOR TRANSDUCTION SYSTEMS.
CC -1- SIMILARITY: TO OTHER PROKARYOTIC SENSOR TRANSDUCTION HISTIDINE
CC KINASES.
DR EMBL: AF006176; AA033464.1;
DR InterPro: IPR000410; Bctrl_sens.
DR InterPro: IPR000658; DUF5.
DR InterPro: IPR003660; HAMF.
DR InterPro: IPR003594; HATPase_C.
DR InterPro: IPR003661; His_kinA.

```

```

DR InterPro: IPR001789; Response_reg.
DR Pfam: PF00672; DUF5. 1.
DR Pfam: PF00072; response_reg. 1.
DR Pfam: PF00512; signal. 1.
DR PRINTS: PR00344; BCTRLSENSOR.
DR SMART: SM00304; HAMF. 1.
DR SMART: SM00387; HATPase_C. 1.
DR SMART: SM00388; HSKA. 1.
DR SMART: SM00448; REC. 1.
KW Complete proteome; Kinase; Phosphorylation; Sensory transduction;
KW Transferase.
SQ SEQUENCE 862 AA; 97846 MW; AB6814A3B63626C0 CRC64;

Query Match
Best Local Similarity 2.0%; Score 9; DB 2; Length 862;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LSTLTSL 13
Db 293 LSTLTSL 301

RESULT 4
O9AX85 ID 09AX85 PRELIMINARY: PRT: 1432 AA.
AC 09AX85:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PUTATIVE ABC TRANSPORTER PROTEIN.
GN P0410E03.7.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartiaceae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-CV. NIPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa niponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone: P0410E03."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP002844; BAB21276.1;
SQ SEQUENCE 1432 AA; 160234 MW; 04E8CF2CA08D993 CRC64;

Query Match
Best Local Similarity 2.0%; Score 9; DB 10; Length 1432;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 FTPLSLFD 178
Db 830 FTPLSLFD 838

RESULT 5
O9M268 ID 09M268 PRELIMINARY: PRT: 299 AA.
AC 09M268:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE RAV-LIKE PROTEIN.
GN F21F14.140.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN 11
RP SEQUENCE FROM N.A.

```


RA Choliene N., Robert C., Brotlier P., Wincker P., Catolico L.,
 RA Artiguenave F., Saurin W., Weissenbach J., Mewes H.W., Lemcke K.,
 RA Mayer K.F.X., Queller F., Salanoubat M.,
 RA Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [12]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL138642; CAB71904.1; -.
 DR InterPro: IPR003340; B3.
 DR Pfam: PF02362; B3; 1.
 SQ SEQUENCE 299 AA; 34270 MW; 2AB841F179DD174B CRC64;

Query Match 1.8%; Score 8; DB 10; Length 299;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 174 SLSPDLDR 181
 DB 292 SLSPDLDR 299

RESULT 6
 ID 09CL22 PRELIMINARY; PRT; 306 AA.
 AC 09CL22;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE HYPOTHETICAL PROTEIN PM1426.
 GN PM1426.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Pasteurella.
 OX NCBL_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PM70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.,
 RT "Complete genomic sequence of Pasteurella multocida Pm70.",
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 DR EMBL: AE006179; AK03510.1; -.
 DR InterPro: IPR003187; PLA1.
 DR Pfam: PF02253; PLA1; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 306 AA; 35580 MW; EAF3DE8C1C22B26E CRC64;

Query Match 1.8%; Score 8; DB 2; Length 306;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 314 KLSRSNMR 321
 DB 184 KLSRSNMR 191

RESULT 7
 ID 067608 PRELIMINARY; PRT; 55 AA.
 AC 067608;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE SUBGENOMIC DNA DERIVED FROM DNA B CCCDS - COVALENTLY CLOSED CIRCULAR
 DE DOUBLE-STRANDED MOLECULE.
 OS Tomato golden mosaic virus (TGMV).
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
 OX NCBL_TaxID=10831;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=87040767; PubMed=3022243;
 RA McDowell S.W., Coutts R.H.A., Buck K.W.,
 RT "Molecular characterisation of subgenomic single-stranded and double-
 RT stranded DNA forms isolated from plants infected with tomato golden
 RT mosaic virus.",
 RL Nucleic Acids Res. 14:7967-7984(1986).
 DR EMBL: X04485; CAA28171.1; -.
 DR InterPro: IPR000211; Gemin1_B1.
 DR Pfam: PF00845; Gemin1_B1; 1.
 SQ SEQUENCE 55 AA; 6281 MW; 41963570D739EC96 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 55;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 312 SAKLSRS 318
 DB 43 SAKLSRS 49

RESULT 8
 ID 091ZS5 PRELIMINARY; PRT; 72 AA.
 AC 091ZS5;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE COAT PROTEIN (FRAGMENT).
 GN CP.
 OS Potato virus Y strain NTN (PVY(NTN)).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Potyvirus.
 OX NCBL_TaxID=122280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TUBER NECROSING (NTN);
 RA Czerovska N., Filigiarova M., Moravec T., Petrzik K.,
 RT "Differences in nucleotide and amino acid sequences of N-terminal
 RT parts of coat proteins among isolates of potato virus Y-NTN strain.",
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF228633; AAF67874.1; -.
 FT NON_TER 1 1
 FT NON_TER 72 72
 SQ SEQUENCE 72 AA; 7721 MW; 2C5287E15624CBDD CRC64;

Query Match 1.6%; Score 7; DB 12; Length 72;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 156 TKKDAKQ 162
 DB 11 TKKDAKQ 17

RESULT 9
 ID 091ZS4 PRELIMINARY; PRT; 72 AA.
 AC 091ZS4;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE COAT PROTEIN (FRAGMENT).
 GN CP.
 OS Potato virus Y strain NTN (PVY(NTN)).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Potyvirus.
 OX NCBL_TaxID=122280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TUBER NECROSING (NTN);
 RA Czerovska N., Filigiarova M., Moravec T., Petrzik K.,

RT "Differences in nucleotide and amino acid sequences of N-terminal
RT parts of coat proteins among isolates of potato virus Y-NTN strain."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF228634; AAF67875.1; -
FT NON_TER 1 1
SQ SEQUENCE 72 AA; 7692 MW; 4F129B870523B482 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 72;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKO 162
DB 11 TKKDAKO 17

RESULT 10

ID 091ZS3 PRELIMINARY; PRT; 95 AA.
AC 091ZS3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY(NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Polyviruses.
OX NCBI_TaxID=122280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUBER NECROSING (NTN);
RA Czerovska N., Filigayova M., Moravec T., Petrzik K.;
RT "Differences in nucleotide and amino acid sequences of N-terminal
RT parts of coat proteins among isolates of potato virus Y-NTN strain."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF228635; AAF67876.1; -
FT NON_TER 1 1
SQ SEQUENCE 95 AA; 10446 MW; 28C06891777C9928 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 95;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKO 162
DB 11 TKKDAKO 17

RESULT 11

ID 091ZS8 PRELIMINARY; PRT; 96 AA.
AC 091ZS8;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y (strain N) (PVY).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Polyviruses.
OX NCBI_TaxID=12219;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NECROSING (N);
RA Czerovska N., Filigayova M., Moravec T., Petrzik K.;
RT "Differences in nucleotide and amino acid sequences of N-terminal
RT parts of coat proteins among isolates of potato virus Y-NTN strain."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF228630; AAF67871.1; -
FT NON_TER 1 1
SQ SEQUENCE 96 AA; 10559 MW; 89C3A5675B9784CA CRC64;

Query Match 1.6%; Score 7; DB 12; Length 96;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKO 162
DB 11 TKKDAKO 17

RESULT 12

ID 012148 PRELIMINARY; PRT; 100 AA.
AC 012148;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE ORF YDL163W.
GN YDL163W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Pohl T.M.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MIPS;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ALPHAS288C;
RA Pohl T.M.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: 274212; CAA98738.1; -
DR EMBL: 267750; CAA91583.1; -
DR SGD: S0002322; YDL163W.
KW Hypothetical protein.
SQ SEQUENCE 100 AA; 10751 MW; 952FB92C852AA710 CRC64;

Query Match 1.6%; Score 7; DB 3; Length 100;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LFLSILS 14
DB 26 LFLSILS 32

RESULT 13

ID 091ZS6 PRELIMINARY; PRT; 101 AA.
AC 091ZS6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY(NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Polyviruses.
OX NCBI_TaxID=122280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUBER NECROSING (NTN);

RA Carovska N., Filigajrova M., Moravec T., Petrzik K.;
 RT "Differences in nucleotide and amino acid sequences of N-terminal
 RL parts of coat proteins among isolates of potato virus Y-NTN strain."
 DR EMBL: AF228632; AAF67873.1; -
 FT NON_TER 1
 FT SEQUENCE 101 AA; 11112 MW; B55BF1C3D5FE9BD CRC64;

Query Match 1.6%; Score 7; DB 12; Length 101;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKQ 162
 |||||
 DB 11 TKKDAKQ 17

RESULT 14
 ID Q9CIE9 PRELIMINARY; PRT; 116 AA.
 AC Q9CIE9;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DE CELLOBIOSE-SPECIFIC PTS SYSTEM IIA COMPONENT (EC 2.7.1.69).
 GN PTCA.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria: Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Lactococcus;
 OX NCBI_TaxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL1403;
 RA Bolotin A., Wincker P., Manger S., Jaillon O., Malarne K.,
 RT Weissensbach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus
 RT lactis."
 RL Genome Res. 0:0-0(2001).
 DR EMBL: AE006278; AAK04512.1; -
 DR InterPro: IPR003188; PTS_IIA.
 DR Pfam: PF02255; PTS_IIA; 1.
 KW Complete proteome.
 SQ SEQUENCE 116 AA; 12814 MW; 995F034417DA3DA0 CRC64;

Query Match 1.6%; Score 7; DB 2; Length 116;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 AIOQAKA 26
 |||||
 DB 34 AIOQAKA 40

RESULT 15
 ID Q9RU16 PRELIMINARY; PRT; 122 AA.
 AC Q9RU16;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DE HYPOTHETICAL 13.7 KDA PROTEIN.
 GN DR1580.
 OS Deinococcus radiodurans.
 OC Bacteria: Thermus/Deinococcus group; Deinococcales; Deinococcus.
 OX NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=R1;
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,

RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome sequence of the radioresistant bacterium Deinococcus
 RT radiodurans R1."
 RL Science 286:1571-1577(1999).
 DR EMBL: AE002002; AAF1150.1; -
 DR TIGR: DR1580; -
 DR InterPro: IPR000325; Glyoxalase_1.
 DR Pfam: PF00903; Glyoxalase; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 122 AA; 13683 MW; 7BA836DCAD786E51 CRC64;

Query Match 1.6%; Score 7; DB 2; Length 122;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 EHEPELY 76
 |||||
 DB 115 EHEPELY 121

Search completed: May 9, 2002, 13:05:18
 Job time: 613 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:54:59 ; Search time 72.68 Seconds
(without alignments)
450.473 Million cell updates/sec

Title: US-09-787-083-8

Perfect score: 442
Sequence: 1 MKVSLSTLTLSTLPCFALLA.....YNHATSPGVGLMDNMGL 442

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 522463 seqs, 74073290 residues

Word size : 0

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

A_Geneseq_1101.*
1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	442	100.0	442	21	BASB034 amino acid
2	327	74.0	442	21	BASB034 amino acid
3	316	71.5	442	21	BASB034 amino acid
4	289	65.4	442	21	BASB034 amino acid
5	9	2.0	370	21	Neisseria gonorrhoe
6	9	2.0	370	21	Neisseria meningit
7	9	2.0	370	21	Neisseria meningit
8	9	2.0	374	21	Neisseria meningit
9	9	2.0	375	21	Neisseria meningit
10	8	1.8	278	21	Arabidopsis thalia
11	8	1.8	283	21	Arabidopsis thalia

12	8	1.8	299	21	AG11135	Arabidopsis thalia
13	7	1.6	16	15	AAR52127	Mouse light chain
14	7	1.6	50	21	AAT9463	Rat alpha tubulin
15	7	1.6	64	21	AG06322	Arabidopsis thalia
16	7	1.6	82	21	AG01488	Human secreted pro
17	7	1.6	106	21	AG07608	Arabidopsis thalia
18	7	1.6	106	21	AG51952	Arabidopsis thalia
19	7	1.6	108	21	AG37811	Arabidopsis thalia
20	7	1.6	116	21	AA187313	Human signal pepti
21	7	1.6	136	21	AG37810	Arabidopsis thalia
22	7	1.6	158	21	AA79176	Haematopoietic ste
23	7	1.6	169	20	AA34661	Chlamydia pneumoni
24	7	1.6	181	21	AG37809	Arabidopsis thalia
25	7	1.6	186	21	AG07607	Arabidopsis thalia
26	7	1.6	186	21	AG51951	Arabidopsis thalia
27	7	1.6	188	21	AG51950	Arabidopsis thalia
28	7	1.6	190	21	AG07606	Arabidopsis thalia
29	7	1.6	268	13	AAR21542	Envelope protein o
30	7	1.6	302	13	AAR24254	Protease inhibitor
31	7	1.6	316	21	AG39499	Arabidopsis thalia
32	7	1.6	324	21	AG39499	Arabidopsis thalia
33	7	1.6	326	13	AAR24356	Protease inhibitor
34	7	1.6	365	21	AG39498	Arabidopsis thalia
35	7	1.6	381	20	AA60241	Human endometrium
36	7	1.6	448	21	AA91960	Human cytoskeleton
37	7	1.6	523	21	AG44285	Arabidopsis thalia
38	7	1.6	552	16	AAR67007	Firefly luciferase
39	7	1.6	666	20	AA21871	Amino acid sequenc
40	7	1.6	683	20	AA21869	Amino acid sequenc
41	7	1.6	684	20	AA21873	Amino acid sequenc
42	7	1.6	715	20	AA21860	Amino acid sequenc
43	7	1.6	783	20	AAW0586	Human SLAP-130. H
44	7	1.6	785	20	AAW0586	Human SLAP-130. H
45	7	1.6	900	20	AAW6264	Human axlin. Homo

ALIGNMENTS

RESULT 1	
AA85271	standard; Protein: 442 AA.
XX	
AC	AA85271:
XX	
DT	29-JUN-2000 (first entry)
XX	
DE	BASB034 amino acid sequence #4.
KW	Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KW	vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW	sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW	hearing loss; antibacterial drug.
XX	
OS	Moraxella catarrhalis.
PN	W020015802-A1.
XX	
PD	23-MAR-2000.
XX	
PF	14-SEP-1999; 99MO-EP06781.
XX	
PR	14-SEP-1998; 98GB-0020002.
XX	
PA	(SMK) SMITHKLINE BEECHAM BIOLOGICALS.
XX	
PI	Ruelle J;
XX	
DR	WPI; 2000-271440/23.
XX	
PT	N-PSDB; AAA10703.
XX	
PT	Novel BASB034 polynucleotides and polypeptides from Moraxella
	catarrhalis used to prepare vaccines against bacterial infections

XX Claim 3; Page 69; 106pp; English
PS

This sentence represents a Moraxella catarrhalis BASB034 polypeptide from strain Mc2869. The invention relates to BASB034 polypeptides from M. catarrhalis strains Mc2931, Mc2508, Mc2913 and Mc2969. The BASB034 polynucleotides and polypeptides may be employed as research reagents and material for the discovery of treatments and diagnostics for diseases, particularly human diseases. They are particularly used to diagnose and treat M. catarrhalis infections. They can be used for diagnosis of disease, staging of disease, or determining response of an infectious organism to drugs. The polynucleotides may be used as a source for hybridization probes, and for screening of genetic mutations, serotype, organism or strain identification, identification of mutations in BASB034 sequences, and as components of arrays which are useful for diagnostic and prognostic purposes. The polypeptides can be used to produce antibodies. The polypeptides can also be used in vaccine formulations, and to identify agonists and antagonists. The polypeptides, antibodies, agonists and antagonists (which are bacteriostatic) are used for the treatment and prevention of diseases such as otitis media in infants and children, pneumonia in elderslies, sinusitis, nosocomial infections and invasive diseases, and chronic otitis media with hearing loss. The polypeptides, agonists and antagonists are also used for screening of antibacterial drugs. The BASB034 products of the invention can be used screen for new antibacterial compounds that may target resistant bacteria.

Sequence 442 AA;

Query Match	100.0%;	Score 442;	DB 21;	Length 442;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 442;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY	1	MKXSLSTLTLSLRFCALALAQQAQVNPRAFAFDEYKRSNDLGDQNELRITDVQATQSA	60
Db	1	mkvslstltlslrpfallaqgaqavnpvrafvdevyrsndlqgdndelrldvqatsqsa	60
QY	61	STD7ANPLDEHSBELYTTALFNKTMNLNCSALNDINRLACUDPLVNGEPAYIKTKRSI	120
Db	61	stdtanpldehsbelrlyttalenktnmlncsalndinrlacudplvngespayiktkrsi	120
QY	121	RLDETIMOTIGKRPQVYQOETTRDFLMGNENKGMLTGKDAQOLEYAAKQFTPLSLSTDD	180
Db	121	rldetimotigkrpqvyqoettrdfilmgnenkmgtgkdaqoleyaaqftplslstdd	180
QY	181	RNNPFIASSRPHNPYUVRITFMHKKRPNSRPTSEHNEKQSTPNBFARDELKQVSYVYKA	240
Db	181	rnnpfiassrphnpuyvritfmhkkrpnsrptsehnekqstpnbfarelkqvsvyuka	240
QY	241	AEDLGTSDSDLMFGYTOOSHMOJINGNKSRRFRYNDQOPELFIPLQPVYSDLPMDGKRYM	300
Db	241	aedlgtssdldmfgytsooshmojingnksrrfryndqoapelfiplqpvysdlpmdgkrym	300
QY	301	GMGAVNHNSNGESAKLSRSWNAAYUIMAGMEKNLTVMRTRIGRFRKESSGQRPDDNPDILD	360
Db	301	gmgaavnhnsngesaklsrswnaayuiimagmeknltvmrtrigrfrkessgqrpddnpdild	360
QY	361	YUGYDVAFFLQOLENKSNSIGTCVRNRPNSGALQDLVYVPLGSGIGSYQIFQFGQXSL	420
Db	361	yugydvafrflqlenksnnsigtcvrnrpnsгалdlvyvplgsgigsyqifqfgqxsll	420
QY	421	IDYNHAEATSPGVGLMLDNMGL	442
Db	421	idyneaatsfgvglmldnmgl	442

RESULT	2
AAV85268	
ID	AAV85268 standard; Protein; 442 AA
XX	
AC	AAV85268;
XX	

DT 29-JUN-2000 (first entry)
XX
DE BASB034 amino acid sequence #1.
XX
KM Moraxella catarrhalis infection; BASB034; diagnosis; straining;
KM vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KM sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KM hearing loss; antibacterial drug.
XX
OS Moraxella catarrhalis.
XX
PN WO200015802-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99WO-EP06781.
XX
PR 14-SEP-1998; 98GB-0020002.
XX
PA (SMRK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI
XX Ruehle J;
XX
XX WPI: 2000-271440/23.
DR N-PSDB: AAA10700.
XX
PT Novel BASB034 polynucleotides and polypeptides from Moraxella
PT catarrhalis used to prepare vaccines against bacterial infections -
XX
PS Claim 3; Fig 2; 106pp; English.

Claim 3; Fig 2; 106pp; English

This sequence represents a Moraxella catarrhalis BASB034 polypeptide from strain MC2931 (ATCC 43617). The invention relates to BASB034 polypeptides from M. catarrhalis strains MC2931, MC2908, MC2913 and MC2969. The BASB034 polynucleotides and polypeptides may be employed as research reagents and material for the discovery of treatments and diagnostics for diseases, particularly human diseases. They are particularly used to diagnose and treat M. catarrhalis infections. They can be used for diagnosis of disease, staging of disease, or determining response of an infectious organism to drugs. The polynucleotides may be used as a source for hybridization probes, and for screening of genetic mutations, serotype, organism or strain identification, identification of mutations in BASB034 sequences, and as components of arrays which are useful for diagnostic and prognostic purposes. The polypeptides can be used to produce antibodies. The polypeptides can also be used in vaccine formulations, and to identify agonists and antagonists. The polypeptides, antibodies, agonists and antagonists (which are bacteriostatic) are used for the treatment and prevention of diseases such as otitis media in infants and children, pneumonia in elderlies, sinusitis, nosocomial infections and invasive diseases, and chronic otitis media with hearing loss. The polypeptides, agonists and antagonists are also used for screening of antibacterial drugs. The BASB034 products of the invention can be used screen for new antibacterial compounds that may target resistant bacteria.

Sequence	442 AA;
5Q	

Query Match	74.0%;	Score 327;	DB 21;	Length 442;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches	427;	Conservative	0;	Matches 1;
			Indels	0;
			Gaps	0;
QY	15	CPAIIAIOAOAVPNPVAFVDEVRSEBNDGQONEPLIDVQATOSASTPANPLDNEHE	74	
DB	15	cfallaiaqgagvnpvafvdevrseendqgnehplidvqastqsaatclanpldnehe	74	
QY	75	LYTTALENKTMLINCSALNODIMRLACYDVLVHGETHPAVIKTRSRIRLDETITQTIKGR	134	
DB	75	lyttalenktmlincsalngdmrlacyctlvbgepvaviketrslrldetltwctikgr	134	
QY	135	QVVYQETTPRPLMGNEKMKLTRKDAKOLEVAAKQTPSLSEFDDLRNRTPLMSSRPNRP	194	
DB	135	qvvlyqettprlmgnekmlctrkdaqlvyaakqfslpslsefdldrnrtplmssrpnrp	194	

QY 195 MYVLPITFHGKPNRSPNTPSHAKOFNPEERAPELKFOVSVKYKAAEDLWGTSDLMWG 254
| | | | |
Db 195 myvlpitfhgkpnrspnspnshakqfnpneirapelkfqvsvkvaadlwgtsdllwfg 254
QY 255 YTGQSHMOIFGKKSRPFVRHDYQPELFTQPVYSDLPWDKVRMIGCAVHNSGESAK 314
| | | | |
Db 255 ytgqshmqifgknsrpfvrhdypelftlqpvysdlpwdgkvrmlgmgaahnsgeesak 314
QY 315 LSRSMNRAVYLAGMEKMKLVMPRIWGRIFKESGSGQPDNDPILIDYGYGVDFRYOLE 374
| | | | |
Db 315 lsrsmnraylagmewkmltvmprlwgriifkessgsqpdndpildyygydvrflyql 374
QY 375 NKSNSGTVRRNPRSGKALQLDVYVPLGKISGTFQIFQGYGQSLIDYNHEATSFQVL 434
| | | | |
Db 375 nksnsgtvrrnprsgkalgldvypylgklsyfqifqgygsllidynheatsfgyvl 434
QY 435 MLNDMGL 442
| | | | |
Db 435 mlndmgl 442

RESULT 3
AAV85270 ID AAV85270 standard; Protein: 442 AA.
XX AC AAV85270;
XX DT 29-JUN-2000 (first entry)
XX DE BASB034 amino acid sequence #3.
XX KW Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KM vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KM sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KM hearing loss; antibacterial drug.
XX OS Moraxella catarrhalis.
XX PN W0200015802-A1.
XX PD 23-MAR-2000.
XX PE 14-SEP-1999; 99MO-EP06781.
XX PR 14-SEP-1998; 98GB-0020002.
XX PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX PI Ruelle J;
XX DR WPI: 2000-271440/23.
XX DR N-PSDB; AAA10702.
XX PT Novel BASB034 polynucleotides and polypeptides from Moraxella
XX catarrhalis used to prepare vaccines against bacterial infections
XX PS Claim 3; Page 68; 106pp; English.
XX
CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2913. The invention relates to BASB034 polypeptides from
CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
CC polynucleotides and polypeptides may be employed as research reagents and
CC material for the discovery of treatments and diagnostics for diseases,
CC particularly human diseases. They are particularly used to diagnose and
CC treat M. catarrhalis infections. They can be used for diagnosis of
CC disease, staging of disease, or determining response of an infectious
CC organism to drugs. The polynucleotides may be used as a source for
CC hybridization probes, and for screening of genetic mutations, serotype,
CC organism or strain identification, identification of mutations in BASB034
CC sequences, and as components of arrays which are useful for diagnostic
CC and prognostic purposes. The polypeptides can be used to produce
CC antibodies. The polypeptides can also be used in vaccine formulations.

CC and to identify agonists and antagonists. The polypeptides, antibodies,
CC agonists and antagonists (which are bacteriostatic) are used for the
CC treatment and prevention of diseases such as otitis media in infants and
CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
CC invasive diseases, and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.
XX
SQ Sequence 442 AA;

Query Match 71.5%; Score 316; DB 21; Length 442;
Best Local Similarity 99.8%; Pred. No. 1,3e-298;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 26 AVPNPVAFVDEVRSENDLQDNEPLPIDVQATQASSTDTANPLDEHEPELYTALENKTM 85
| | | | |
Db 26 avpnvpafvdevrseendlqdnelpidvqatqasctdanpldehepelytalenktm 85
QY 86 LINCASALNQDITMLACYDTLVHGETPAVTKRSIRLDETITWQTIKGRPOVYQETDPI 145
| | | | |
Db 86 lincsalnqdlmlacydtllvhgetpavtkrsirlidetlwtqikgkpyvygettdpi 145
QY 146 FLMGNEKMLTKDKAKOLEFAAKOFTPLSLSFDDLDRNTPLMSSRPHNPMVLPITFPHGK 205
| | | | |
Db 146 flmgnekmltkdkakolefayakqftplslsfddlrrnptlmssrphnmpvlpitfphgk 205
QY 206 PNRSPNTPSHAKOFTPEFRAPELKFOVSVKYKAAEDLWGTSDLMFGYQOSHMOIFN 265
| | | | |
Db 206 pnrspnspshakoftpefrapelkfqvsvkvaadlwgtsdllwfgyqgshmqifn 265
QY 266 GKNSRPFVRHDYQPELFTQPVYSDLPWDKVRMIGCAVHNSGESAKLSRSNNRAVLM 325
| | | | |
Db 266 gknsrpfvrhdypelftlqpvysdlpwdgkvrmlgmgaahnsgeesaklsrsnraylm 325
QY 326 AGMEKKNLTVMPRIWGRIFKESGSGQPDNDPILIDYGYGVDFRYOLENKSNSITGVRY 385
| | | | |
Db 326 agmekknltpmriwgrifkessgsqpdndpildyygydvrflyqlenksnsgtvry 385
QY 386 NPSRGKALQLDVYVPLGKISGTFQIFQGYGQSLIDYNHEATSFQVGLMLNDMGL 442
| | | | |
Db 386 npsrgkalgldvypylgklsyfqifqgygsllidynheatsfgyvlmlndmgl 442

RESULT 4
AAV85269 ID AAV85269 standard; Protein: 442 AA.
XX AC AAV85269;
XX DT 29-JUN-2000 (first entry)
XX DE BASB034 amino acid sequence #2.
XX KW Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KM vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KM sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KM hearing loss; antibacterial drug.
XX OS Moraxella catarrhalis.
XX PN W0200015802-A1.
XX PD 23-MAR-2000.
XX PE 14-SEP-1999; 99MO-EP06781.
XX PR 14-SEP-1998; 98GB-0020002.
XX PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX

PI Ruehle J.
XX
DR WPI: 2000-271440/23.
DR N-PSDB: AAA10701.
XX
PT Novel BAsB034 polynucleotides and polypeptides from *Moraxella*
PT catarrhalis used to prepare vaccines against bacterial infections -
XX
PS Claim 3; Page 67; 106pp: English.

Query Match	65.4%	Score 289	DB 21	Length 442
Best Local Similarity	99.7%	Pred. No. 2.3e-272		
Matches 389	Conservative	0	Mismatches 1	Indels 0
				Gaps 0

0Y	53	VOSATQASSTPTANPLDNDHEBELXTTALLKENTMYLINSALNODIMRLACVDTLVHGSEPA	11.2
Dd	53	vsaatqassttclanpldndehelxtallkentakmlinsalngdlmrtaacydclvhgscpa	11.2
0Y	113	VIKTKRSrRLDEETIMQTIKGRQOVVYOETTDPIFLMNGEKMLKKPAKOLEYAAKQFTP	17.2
Dd	113	vIktrksrlldetlmtqtlkgrpqvvyyetctdplflmgnekmyllckakqleyaakqftp	17.2
0Y	173	LSLSFDLDRNNTPLMSSRPNNPKVYLPIFMHGKRNRPNTPSHEAKQFTPNERRAPLKF	23.2
Dd	173	lslsfldlrnntplmssrphnnpmyvylpflmhgkpnrsnpshsraqftpnefrapelkf	23.2
0Y	233	QVSYVKAEDLMGDSOJLMRGYQOOSIMOJFNKNSRPPRVNHJOPEFLTOVYSDLP	23.2
Dd	233	qvsyvukaedlmgdsdjlwfrgyqqsjwqjflngksnrpfvnhdyqeflfcqpyvsdlp	23.2
0Y	293	WDGKRYRtGMAVHNSNEESAKLSRSWNRRALtLMAEMKMLTTPVAPRtMGRtIFKESGSSOP	35.2
Dd	293	wdgkryrimgavvhnsngesaklsrswrrlalmgmekmlttmprrtlmgjlfksgssqp	35.2
0Y	353	DDNPDIIDLYtYGtGVDRFLYOLENKSNTISGTVRYNPRSGKALQJLDLYYPLGKtISGtEQI	41.2
Dd	353	ddnpdlldyugdyrflygleknsnlsgtvyrrpsgkqalqldyyplgktsygyfqi	41.2
0Y	413	FOGYGOSLIDYVNHAEATSGVGLMINDMGL	44.2
Dd	413	fgygygsldyvhaeatstgvgldmndmgll	44.2

RESULT	5
AAV75156	
ID	AAV75156 standard; Protein; 370 AA.

XX	AA75156;	
AC		
XX		
DT	21-MAR-2000 (first entry)	
XX		
DE	Neisseria gonorrhoeae ORF 582 protein sequence SEQ ID NO:1786.	
XX		
KW	Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;	
KW	antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;	
XX	antibacterial; gene therapy.	
XX		
OS	Neisseria gonorrhoeae.	
XX		
PN	W09957280-A2.	
XX		
PD	11-NOV-1999.	
XX		
PF	30-APR-1999; 99MO-US09346.	
XX		
PR	01-MAY-1998; 98US-0083758.	
PR	31-JUL-1998; 98US-0094869.	
PR	02-SEP-1998; 98US-0098994.	
PR	02-SEP-1998; 98US-0099062.	
PR	09-OCT-1998; 98US-0103749.	
PR	09-OCT-1998; 98US-0103794.	
PR	09-OCT-1998; 98US-0103796.	
PR	25-FEB-1999; 99US-0121528.	
XX		
PA	(CHIR) CHIRON CORP.	
PA	(GENO-) INST GENOMIC RES.	
XX		
PI	Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;	
PI	Petersen J, Piza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;	
PI	Tetzelin H, Venter JC.	
XX		
DR	WPI: 2000-062150/05.	
DR	N-PSDB; AA253918.	
XX		
PT	Novel Neisserial polypeptides predicted to be useful antigens for vaccines and diagnostics -	

```

Query Match      2.0%; Score 9; DB 21; Length 370;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 279 PEIPLTPV 287
      ||| |||||
Db 209 PEIPLTPV 217

RESULT 6
AA75157
ID AA75157 standard; Protein: 370 AA.
XX

```



```

AC  AAY75157;
XX
XX  21-MAR-2000 (first entry)
XX
XX  Neisseria meningitidis ORF 582 protein sequence SEQ ID NO:1788.
DE
XX  Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
XX  antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
XX  antibacterial; gene therapy.
XX
XX  Neisseria meningitidis.
OS
XX  WO957280-A2.
XX
XX  11-NOV-1999.
XX
XX  30-APR-1999; 99WO-US09346.
XX
XX  01-MAY-1998; 98US-0083758.
XX  31-JUL-1998; 98US-0094869.
XX  02-SEP-1998; 98US-0098994.
XX  02-SEP-1998; 98US-0099062.
XX  09-OCT-1998; 98US-0103749.
XX  09-OCT-1998; 98US-0103794.
XX  09-OCT-1998; 98US-0103796.
XX  25-FEB-1999; 99US-0121528.
XX
XX  (CHIR ) CHIRON CORP.
XX  (GENO-) INST GENOMIC RES.
XX
XX  Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;
XX  Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
XX  Tettelin H, Venter JC;
XX
XX  WPI: 2000-062150/05.
XX  N-PSDB; AA253919.
XX
XX  Novel Neisserial polypeptides predicted to be useful antigens for
XX  vaccines and diagnostics -
XX
XX  Claim 2; Page 904; 1453pp; English.
XX
XX  AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
XX  represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
XX  and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
XX  PCR primers used in the exemplification of the present invention. The
XX  polypeptides, the polynucleotides, antibodies and compositions of
XX  the invention can be used as vaccines, as diagnostic reagents, and as
XX  immunogenic compositions. The polypeptides can be used in the
XX  manufacture of medicaments for treating or preventing infection due to
XX  Neisserial bacteria (e.g. meningitis and septicemia), to detect the
XX  presence of Neisseria bacteria, or to raise antibodies. They may also
XX  be used to screen for agonists or antagonists, which may themselves
XX  have use as antibacterial agents. The polynucleotides of the invention
XX  may also be used in gene therapy protocols.
XX
XX
XX  Sequence 370 AA;
XX
XX  Query Match 2.0%; Score 9; DB 21; Length 370;
XX  Best Local Similarity 100.0%; Pred. No. 2.1;
XX  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX  QY 279 PEIFLTQPV 287
XX  |||||||
XX  Db 209 pelfltqpv 217

```

```

XX
XX  21-MAR-2000 (first entry)
XX
XX  Neisseria meningitidis ORF 582 protein sequence SEQ ID NO:1790.
DE
XX  Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
XX  antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
XX  antibacterial; gene therapy.
XX
XX  Neisseria meningitidis.
OS
XX  WO957280-A2.
XX
XX  11-NOV-1999.
XX
XX  30-APR-1999; 99WO-US09346.
XX
XX  01-MAY-1998; 98US-0083758.
XX  31-JUL-1998; 98US-0094869.
XX  02-SEP-1998; 98US-0098994.
XX  02-SEP-1998; 98US-0099062.
XX  09-OCT-1998; 98US-0103749.
XX  09-OCT-1998; 98US-0103794.
XX  09-OCT-1998; 98US-0103796.
XX  25-FEB-1999; 99US-0121528.
XX
XX  (CHIR ) CHIRON CORP.
XX  (GENO-) INST GENOMIC RES.
XX
XX  Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;
XX  Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
XX  Tettelin H, Venter JC;
XX
XX  WPI: 2000-062150/05.
XX  N-PSDB; AA253920.
XX
XX  Novel Neisserial polypeptides predicted to be useful antigens for
XX  vaccines and diagnostics -
XX
XX  Claim 2; Page 905; 1453pp; English.
XX
XX  AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
XX  represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
XX  and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
XX  PCR primers used in the exemplification of the present invention. The
XX  polypeptides, the polynucleotides, antibodies and compositions of
XX  the invention can be used as vaccines, as diagnostic reagents, and as
XX  immunogenic compositions. The polypeptides can be used in the
XX  manufacture of medicaments for treating or preventing infection due to
XX  Neisserial bacteria (e.g. meningitis and septicemia), to detect the
XX  presence of Neisseria bacteria, or to raise antibodies. They may also
XX  be used to screen for agonists or antagonists, which may themselves
XX  have use as antibacterial agents. The polynucleotides of the invention
XX  may also be used in gene therapy protocols.
XX
XX
XX  Sequence 370 AA;
XX
XX  Query Match 2.0%; Score 9; DB 21; Length 370;
XX  Best Local Similarity 100.0%; Pred. No. 2.1;
XX  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX  QY 279 PEIFLTQPV 287
XX  |||||||
XX  Db 209 pelfltqpv 217

```

```

RESULT 7
AAY75158
ID AAY75158 standard; Protein: 370 AA.
XX
XX  AAY75158;
AC

```

```

RESULT 8
AAY70629
ID AAY70629 standard; Protein: 374 AA.
XX
XX  AAY70629;
AC

```

DT 18-JUL-2000 (first entry)
 XX Neisseria meningitidis serogroup B strain H44/76 BASB033 protein.
 DE BASB033: diagnosis, prophylaxis; treatment; antibacterial; vaccine;
 XX Neisseria meningitidis infection.
 KM Neisseria meningitidis infection.
 OS Neisseria meningitidis.
 XX WO200015801-A1.
 PN 23-MAR-2000.
 PD 09-SEP-1999; 99WO-EP06718.
 XX 14-SEP-1998; 98GB-0020003.
 PR (SMK) SMITHKLINE BECHAM BIOLOGICALS.
 PA Ruelle J;
 XX WPI: 2000-271439/23.
 DR N-PSDB; AA252134.
 XX Isolated BASB033 polypeptides and polynucleotides of Neisseria
 PT meningitidis, useful for diagnosis, prophylaxis and treatment of N.
 PS meningitidis infection -
 XX Claim 4; Page 59; 93pp; English.
 CC The present sequence is a BASB033 protein from
 CC Neisseria meningitidis serogroup B strain H44/76. The protein
 CC shows homology to the Klebsiella pneumoniae outer membrane
 CC phospholipase A. The present sequence is useful for diagnosis,
 CC prophylaxis and treatment of N. meningitidis infection. It may also be
 CC used for the discovery and development of antibacterial compounds and
 CC in vaccine compositions.
 CC Sequence 374 AA;
 SQ

Query Match 2.0%; Score 9; DB 21; Length 374;
 Best Local Similarity 100.0%; Pred. No. 2.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 279 PEIFLTQPV 287
 |||||||||
 Db 213 PEIFLTQPV 221

RESULT 9
 AAY70628 standard; Protein; 375 AA.
 ID AAY70628;
 AC AAY70628;
 XX 18-JUL-2000 (first entry)
 DT Neisseria meningitidis serogroup B strain ATCC13090 BASB033 protein.
 DE Neisseria meningitidis infection.
 XX BASB033: diagnosis, prophylaxis; treatment; antibacterial; vaccine;
 KM Neisseria meningitidis infection.
 OS Neisseria meningitidis.
 XX WO200015801-A1.
 PN 23-MAR-2000.
 PD 09-SEP-1999; 99WO-EP06718.
 XX 14-SEP-1998; 98GB-0020003.
 PR

PA (SMK) SMITHKLINE BECHAM BIOLOGICALS.
 XX Ruelle J;
 PI WPI: 2000-271439/23.
 DR N-PSDB; AA252133.
 XX Isolated BASB033 polypeptides and polynucleotides of Neisseria
 PT meningitidis, useful for diagnosis, prophylaxis and treatment of N.
 PS meningitidis infection -
 XX Claim 4; Page 58; 93pp; English.
 CC The present sequence is a BASB033 protein from
 CC Neisseria meningitidis serogroup B strain ATCC13090. The protein
 CC shows homology to the Klebsiella pneumoniae outer membrane
 CC phospholipase A. The present sequence is useful for diagnosis,
 CC prophylaxis and treatment of N. meningitidis infection. It may also be
 CC used for the discovery and development of antibacterial compounds and
 CC in vaccine compositions.
 CC Sequence 375 AA;
 SQ

Query Match 2.0%; Score 9; DB 21; Length 375;
 Best Local Similarity 100.0%; Pred. No. 2.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 279 PEIFLTQPV 287
 |||||||||
 Db 214 PEIFLTQPV 222

RESULT 10
 AAG1137 standard; Protein; 278 AA.
 ID AAG1137;
 AC AAG1137;
 XX 17-OCT-2000 (first entry)
 DT Arabidopsis thaliana protein fragment SEQ ID NO: 9741.
 DE Protein identification; signal transduction pathway; metabolic pathway;
 XX hybridisation assay; genetic mapping; gene expression control; promoter;
 KM termination sequence.
 OS Arabidopsis thaliana.
 XX EPI033405-A2.
 PN 06-SEP-2000.
 PD 25-FEB-2000; 2000EP-0301439.
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130691.
 PR 30-APR-1999; 99US-0131449.
 PR 04-MAY-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999;	99US-0132485.
PR 06-MAY-1999;	99US-0132486.
PR 06-MAY-1999;	99US-0132487.
PR 07-MAY-1999;	99US-0132863.
PR 11-MAY-1999;	99US-0134256.
PR 14-MAY-1999;	99US-0134218.
PR 14-MAY-1999;	99US-0134219.
PR 14-MAY-1999;	99US-0134221.
PR 14-MAY-1999;	99US-0134370.
PR 18-MAY-1999;	99US-0134768.
PR 19-MAY-1999;	99US-0134941.
PR 20-MAY-1999;	99US-0135124.
PR 21-MAY-1999;	99US-0135353.
PR 24-MAY-1999;	99US-0135629.
PR 25-MAY-1999;	99US-0136021.
PR 27-MAY-1999;	99US-0136392.
PR 28-MAY-1999;	99US-0136782.
PR 01-JUN-1999;	99US-0137222.
PR 03-JUN-1999;	99US-0137528.
PR 04-JUN-1999;	99US-0137502.
PR 07-JUN-1999;	99US-0137724.
PR 08-JUN-1999;	99US-0138094.
PR 10-JUN-1999;	99US-0138540.
PR 10-JUN-1999;	99US-0138847.
PR 14-JUN-1999;	99US-0139119.
PR 16-JUN-1999;	99US-0139452.
PR 16-JUN-1999;	99US-0139453.
PR 17-JUN-1999;	99US-0139492.
PR 18-JUN-1999;	99US-0139454.
PR 18-JUN-1999;	99US-0139455.
PR 18-JUN-1999;	99US-0139456.
PR 18-JUN-1999;	99US-0139457.
PR 18-JUN-1999;	99US-0139458.
PR 18-JUN-1999;	99US-0139459.
PR 18-JUN-1999;	99US-0139460.
PR 18-JUN-1999;	99US-0139461.
PR 18-JUN-1999;	99US-0139462.
PR 18-JUN-1999;	99US-0139463.
PR 18-JUN-1999;	99US-0139750.
PR 18-JUN-1999;	99US-0139763.
PR 21-JUN-1999;	99US-0139817.
PR 22-JUN-1999;	99US-0139899.
PR 23-JUN-1999;	99US-0140353.
PR 23-JUN-1999;	99US-0140354.
PR 24-JUN-1999;	99US-0140695.
PR 28-JUN-1999;	99US-0140823.
PR 29-JUN-1999;	99US-0140991.
PR 30-JUN-1999;	99US-0141287.
PR 01-JUL-1999;	99US-0141847.
PR 01-JUL-1999;	99US-0142154.
PR 02-JUL-1999;	99US-0142055.
PR 06-JUL-1999;	99US-0142390.
PR 08-JUL-1999;	99US-0142803.
PR 09-JUL-1999;	99US-0142920.
PR 12-JUL-1999;	99US-0142977.
PR 13-JUL-1999;	99US-0143547.
PR 14-JUL-1999;	99US-0143624.
PR 15-JUL-1999;	99US-0144005.
PR 16-JUL-1999;	99US-0144085.
PR 16-JUL-1999;	99US-0144086.
PR 19-JUL-1999;	99US-0144325.
PR 19-JUL-1999;	99US-0144331.
PR 19-JUL-1999;	99US-0144333.
PR 19-JUL-1999;	99US-0144334.
PR 19-JUL-1999;	99US-0144335.
PR 20-JUL-1999;	99US-0144357.
PR 20-JUL-1999;	99US-0144632.
PR 20-JUL-1999;	99US-0144634.
PR 21-JUL-1999;	99US-0144814.
PR 21-JUL-1999;	99US-0145086.
PR 21-JUL-1999;	99US-0145088.
PR 22-JUL-1999;	99US-0145085.

PR 22-JUL-1999;	99US-0145087.
PR 22-JUL-1999;	99US-0145089.
PR 22-JUL-1999;	99US-0145192.
PR 23-JUL-1999;	99US-0145145.
PR 23-JUL-1999;	99US-0145218.
PR 23-JUL-1999;	99US-0145224.
PR 26-JUL-1999;	99US-0145276.
PR 27-JUL-1999;	99US-0145913.
PR 27-JUL-1999;	99US-0145918.
PR 27-JUL-1999;	99US-0145919.
PR 28-JUL-1999;	99US-0145951.
PR 02-AUG-1999;	99US-0146386.
PR 02-AUG-1999;	99US-0146388.
PR 02-AUG-1999;	99US-0146389.
PR 03-AUG-1999;	99US-0147038.
PR 04-AUG-1999;	99US-0147204.
PR 04-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147192.
PR 05-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 06-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147493.
PR 09-AUG-1999;	99US-0147935.
PR 10-AUG-1999;	99US-0148171.
PR 11-AUG-1999;	99US-0148319.
PR 12-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 13-AUG-1999;	99US-0148684.
PR 16-AUG-1999;	99US-0149368.
PR 17-AUG-1999;	99US-0149175.
PR 18-AUG-1999;	99US-0149426.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 20-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 23-AUG-1999;	99US-0149930.
PR 25-AUG-1999;	99US-0150566.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151080.
PR 30-AUG-1999;	99US-0151303.
PR 31-AUG-1999;	99US-0151438.
PR 01-SEP-1999;	99US-0151930.
PR 07-SEP-1999;	99US-0152363.
PR 10-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155436.
PR 24-SEP-1999;	99US-0155659.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156596.
PR 04-OCT-1999;	99US-0157117.
PR 05-OCT-1999;	99US-0157173.
PR 06-OCT-1999;	99US-0157865.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0159293.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159295.
PR 14-OCT-1999;	99US-0159329.
PR 14-OCT-1999;	99US-0159330.
PR 14-OCT-1999;	99US-0159331.
PR 14-OCT-1999;	99US-0159637.
PR 14-OCT-1999;	99US-0159638.
PR 18-OCT-1999;	99US-0159584.
PR 21-OCT-1999;	99US-0160741.
PR 21-OCT-1999;	99US-0160767.
PR 21-OCT-1999;	99US-0160768.

PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 22-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 8; DB 21; Length 278;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 174 SLSEFLDR 181
Db 271 slsfdldr 278

RESULT 11
AAG1136
ID AAG1136 standard; Protein; 283 AA.
XX AAG1136;
AC
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 9740.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
PD
XX 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159395.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.

PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 8; DB 21; Length 283;
 Best Local Similarity 100.0%; Pred. No. 16;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 SLSFDLDR 181
 |||||
 Db 276 slsfdldr 283

RESULT 12
 AAG11135

ID AAG11135 standard; Protein; 299 AA.

XX AC AAG11135;

XX DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 9739.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
 hybridisation assay; genetic mapping; gene expression control; promoter;
 termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 07-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157553.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.

PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 8; DB 21; Length 299;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 174 SLSFDLDR 181
 |||||
 Db 292 slsfdldr 299

RESULT 13
 AAR52127
 ID AAR52127 standard; Peptide; 16 AA.
 XX AC AAR52127;
 XX DT 27-SEP-1996 (first entry)
 XX DE Mouse light chain surface patch S03410.
 XX KW antibody; humanised; murine; human; heavy chain; light; variable;
 XX KW framework region; complementarity determining region; reshaping;
 XX KW modelling; surface residue; modify.
 XX OS Mus sp.
 XX PN EP592106-A1.
 XX PD 13-APR-1994.
 XX PF 07-SEP-1993; 93EP-0307051.
 XX PR 09-SEP-1992; 92US-0942245.
 XX PA (PEDE/) PEDERSEN J T.
 XX PA (IMMU-) IMMUNOGEN INC.
 XX PI Guild BC, Pedersen JT, Rees AR, Roguska MA, Searle SMJ;
 XX DR WPI; 1994-120230/15.
 XX PS Method of resurfacing of rodent antibodies to produce humanised
 XX PT antibody forms - for producing non-human antibodies with improved
 XX PT therapeutic efficiency by presenting human surface on V-region
 XX PS Example 1; Page 14; 230pp; English.
 XX CC Modification of a rodent antibody or fragment by resurfacing in order
 XX CC to produce a humanised rodent antibody can be determined by calculating
 XX CC homology between murine and human antibody surfaces. In order to test
 XX CC the resurfacing approach of the invention, three humanisation
 XX CC experiments were set up. (1) traditional loop grafting; (2) resurfacing
 XX CC approach using most similar chain; and (3) resurfacing approach using
 XX CC human sequences with most similar surface residues. AAR52069-159 are the
 XX CC surface residue patterns in mouse light chain antibody variable regions.
 XX CC These "patches" were used in the third method, where rodent light and
 XX CC heavy chains were matched and the most similar human sequence found
 XX CC independently only over the surface residues indicated in AAR52030-67.
 XX SQ Sequence 16 AA;

Query Match 1.6%; Score 7; DB 15; Length 16;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 387 PRSGKGA 393
 |||||
 Db 4 prsgkga 10

RESULT 14
 AAY79483
 ID AAY79483 standard; peptide; 50 AA.
 XX AC AAY79483;
 XX DT 01-AUG-2000 (first entry)
 XX DE Rat alpha tubulin variant peptide (aa314-363).
 XX KW Rat; alpha tubulin; mutant; mutein; variant;
 XX KW metabolic x syndrome; hypertension; stroke; diabetes;
 XX KW insulin resistance; obesity; dyslipidemia; anorectic;
 XX KW hypotensive; cerebroprotective; vasotropic; antidiabetic;
 XX KW antilipemic; gene therapy; spontaneously hypertensive; SHR.
 XX OS Rattus sp.
 XX FH Key Location/Qualifiers
 XX FT Peptide 1..50
 XX FT /note= "corresponds to residues 314-363 of
 XX FT full-length protein"
 XX FT Misc-difference 27
 XX FT /note= "replaces Thr in wild-type"
 XX PN WO200018918-A2.
 XX PD 06-APR-2000.
 XX PF 28-SEP-1999; 99WO-US22494.
 XX PR 28-SEP-1998; 98US-0161939.
 XX PA (CURA-) CURAGEN CORP.
 XX PI Shimkets RA;
 XX DR WPI; 2000-303450/26.
 XX PT Novel genes encoding e.g. CD36, SGLT2, and kynurenine aminotransferase,
 XX PT which are predictive and therapeutic for stroke, hypertension, diabetes
 XX PT and obesity -
 XX PS Claim 2; Fig 4G; 79pp; English.
 XX CC This sequence represents amino acids 314-363 of alpha-tubulin
 XX CC from a spontaneously hypertensive rat (SHR). This region of
 XX CC alpha-tubulin carries a Ser for Thr amino acid substitution when
 XX CC compared to the corresponding region in control Wistar Kyoto (WKY)
 XX CC rats (see AAY79484). The alpha-tubulin gene is differentially
 XX CC expressed in SHR, SHR-stroke prone (SP) and WKY rats. It may be
 XX CC involved in a predisposition to vascular injury. The SHR rat
 XX CC is an animal model for human metabolic x syndrome. Genes encoding
 XX CC sodium dependent glucose cotransporter, kynurenine aminotransferase,
 XX CC CD36, aldolase A, atrial natriuretic factor, alpha-cardiac myosin
 XX CC and alpha-tubulin were identified as being potentially associated
 XX CC with hypertension, obesity and insulin resistance. These
 XX CC proteins, including isolated proteins comprising the present
 XX CC sequence, are used for treating, preventing and diagnosing
 XX CC ischemic and metabolic diseases and disorders, such as stroke,
 XX CC hypertension, diabetes and obesity, especially insulin resistivity,
 XX CC dyslipidemia and ischemic stroke (all claimed). Polynucleotides
 XX CC encoding them may also be used in gene therapy and antisense
 XX CC therapy protocols.
 XX SQ Sequence 50 AA;

Query Match 1.6%; Score 7; DB 21; Length 50;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 114 IKTKRSI 120

```
Db      22 iktkrs1 28      |||||
RESULT 15
AAG06322
ID      AAG06322 standard; Protein; 64 AA.
XX
AC      AAG06322;
XX
DT      17-OCT-2000 (first entry)
XX
DE      Arabidopsis thaliana protein fragment SEQ ID NO: 3055.
XX
KW      Protein identification; signal transduction pathway; metabolic pathway;
KW      hybridisation assay; genetic mapping; gene expression control; promoter;
KW      termination sequence.
XX
OS      Arabidopsis thaliana.
PN      EP1033405-A2.
PD
PF      06-SEP-2000.
XX
PF      25-FEB-2000; 2000EP-0301439.
XX
PR      25-FEB-1999; 99US-0121825.
PR      05-MAR-1999; 99US-0123180.
PR      09-MAR-1999; 99US-0123548.
PR      23-MAR-1999; 99US-0125788.
PR      25-MAR-1999; 99US-0126264.
PR      29-MAR-1999; 99US-0126785.
PR      01-APR-1999; 99US-0127462.
PR      06-APR-1999; 99US-0128234.
PR      08-APR-1999; 99US-0128714.
PR      16-APR-1999; 99US-0129845.
PR      19-APR-1999; 99US-0130077.
PR      21-APR-1999; 99US-0130449.
PR      23-APR-1999; 99US-0130510.
PR      28-APR-1999; 99US-0130891.
PR      30-APR-1999; 99US-0131449.
PR      30-APR-1999; 99US-0132048.
PR      04-MAY-1999; 99US-0132407.
PR      05-MAY-1999; 99US-0132484.
PR      06-MAY-1999; 99US-0132485.
PR      06-MAY-1999; 99US-0132486.
PR      07-MAY-1999; 99US-0132863.
PR      11-MAY-1999; 99US-0134256.
PR      14-MAY-1999; 99US-0134218.
PR      14-MAY-1999; 99US-0134219.
PR      14-MAY-1999; 99US-0134221.
PR      14-MAY-1999; 99US-0134370.
PR      18-MAY-1999; 99US-0134768.
PR      19-MAY-1999; 99US-0134941.
PR      20-MAY-1999; 99US-0135124.
PR      21-MAY-1999; 99US-0135353.
PR      24-MAY-1999; 99US-0135629.
PR      25-MAY-1999; 99US-0136021.
PR      27-MAY-1999; 99US-0136392.
PR      28-MAY-1999; 99US-0136782.
PR      01-JUN-1999; 99US-0137222.
PR      03-JUN-1999; 99US-0137528.
PR      04-JUN-1999; 99US-0137502.
PR      07-JUN-1999; 99US-0137724.
PR      08-JUN-1999; 99US-0138094.
PR      10-JUN-1999; 99US-0138540.
PR      10-JUN-1999; 99US-0138847.
PR      14-JUN-1999; 99US-0139119.
PR      16-JUN-1999; 99US-0139452.
PR      16-JUN-1999; 99US-0139453.
PR      17-JUN-1999; 99US-0139492.
PR      18-JUN-1999; 99US-0139454.
PR      18-JUN-1999; 99US-0139455.
PR      18-JUN-1999; 99US-0139456.
PR      18-JUN-1999; 99US-0139457.
PR      18-JUN-1999; 99US-0139458.
PR      18-JUN-1999; 99US-0139459.
PR      18-JUN-1999; 99US-0139460.
PR      18-JUN-1999; 99US-0139461.
PR      18-JUN-1999; 99US-0139462.
PR      18-JUN-1999; 99US-0139463.
PR      18-JUN-1999; 99US-0139750.
PR      18-JUN-1999; 99US-0139763.
PR      21-JUN-1999; 99US-0139817.
PR      22-JUN-1999; 99US-0139899.
PR      23-JUN-1999; 99US-0140353.
PR      23-JUN-1999; 99US-0140354.
PR      24-JUN-1999; 99US-0140695.
PR      28-JUN-1999; 99US-0140823.
PR      29-JUN-1999; 99US-0140991.
PR      30-JUN-1999; 99US-0141287.
PR      01-JUL-1999; 99US-0141842.
PR      01-JUL-1999; 99US-0142154.
PR      02-JUL-1999; 99US-0142055.
PR      06-JUL-1999; 99US-0142390.
PR      08-JUL-1999; 99US-0142803.
PR      09-JUL-1999; 99US-0142920.
PR      12-JUL-1999; 99US-0142977.
PR      13-JUL-1999; 99US-0143542.
PR      14-JUL-1999; 99US-0143624.
PR      15-JUL-1999; 99US-0144005.
PR      16-JUL-1999; 99US-0144085.
PR      16-JUL-1999; 99US-0144086.
PR      19-JUL-1999; 99US-0144325.
PR      19-JUL-1999; 99US-0144331.
PR      19-JUL-1999; 99US-0144332.
PR      19-JUL-1999; 99US-0144333.
PR      19-JUL-1999; 99US-0144334.
PR      20-JUL-1999; 99US-0144335.
PR      20-JUL-1999; 99US-0144352.
PR      20-JUL-1999; 99US-0144632.
PR      20-JUL-1999; 99US-0144884.
PR      21-JUL-1999; 99US-0144814.
PR      21-JUL-1999; 99US-0145086.
PR      21-JUL-1999; 99US-0145088.
PR      22-JUL-1999; 99US-0145085.
PR      22-JUL-1999; 99US-0145087.
PR      22-JUL-1999; 99US-0145089.
PR      22-JUL-1999; 99US-0145192.
PR      23-JUL-1999; 99US-0145145.
PR      23-JUL-1999; 99US-0145218.
PR      23-JUL-1999; 99US-0145224.
PR      26-JUL-1999; 99US-0145276.
PR      27-JUL-1999; 99US-0145913.
PR      27-JUL-1999; 99US-0145918.
PR      27-JUL-1999; 99US-0145919.
PR      28-JUL-1999; 99US-0145951.
PR      02-AUG-1999; 99US-0146386.
PR      02-AUG-1999; 99US-0146388.
PR      02-AUG-1999; 99US-0146389.
PR      03-AUG-1999; 99US-0147038.
PR      04-AUG-1999; 99US-0147204.
PR      04-AUG-1999; 99US-0147302.
PR      05-AUG-1999; 99US-0147192.
PR      05-AUG-1999; 99US-0147260.
PR      06-AUG-1999; 99US-0147303.
PR      06-AUG-1999; 99US-0147416.
PR      09-AUG-1999; 99US-0147493.
PR      09-AUG-1999; 99US-0147935.
PR      10-AUG-1999; 99US-0148171.
PR      11-AUG-1999; 99US-0148319.
PR      12-AUG-1999; 99US-0148341.
PR      13-AUG-1999; 99US-0148565.
PR      13-AUG-1999; 99US-0148684.
PR      16-AUG-1999; 99US-0149368.
```



```
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.6%; Score 7; DB 21; Length 64;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 162 QLEYAAK 168
Db 44 qleyaak 50
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:46:30 ; Search time 36.87 Seconds
(without alignments)
269.771 Million cell updates/sec

Title: US-09-787-083-2
Perfect score: 442
Sequence: 1 MKVSLSTLTLSCFAILA.....YNHEATSGVGLMLNDWGL 442

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 212252 seqs, 22503292 residues

Word size : 0
Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	DB ID	Description
1	7	1.6	16	Sequence 98, Appl
2	7	1.6	347	Sequence 14, Appl
3	7	1.6	552	Sequence 6, Appl
4	7	1.6	900	Sequence 4, Appl
5	7	1.6	934	Sequence 80, Appl
6	7	1.6	1026	Sequence 6, Appl
7	7	1.6	1026	Sequence 6, Appl
8	6	1.4	10	Sequence 30, Appl
9	6	1.4	11	Sequence 29, Appl
10	6	1.4	12	Sequence 28, Appl
11	6	1.4	13	Sequence 27, Appl
12	6	1.4	14	Sequence 27, Appl
13	6	1.4	14	Sequence 30, Appl
14	6	1.4	14	Sequence 28, Appl
15	6	1.4	14	Sequence 28, Appl
16	6	1.4	15	Sequence 5, Appl
17	6	1.4	16	Sequence 4, Appl
18	6	1.4	20	Sequence 1, Appl
19	6	1.4	21	Sequence 13, Appl
20	6	1.4	21	Sequence 15, Appl
21	6	1.4	21	Sequence 1, Appl
22	6	1.4	21	Sequence 1, Appl
23	6	1.4	21	Sequence 13, Appl
24	6	1.4	21	Sequence 15, Appl
25	6	1.4	21	Sequence 1, Appl
26	6	1.4	21	Sequence 1, Appl
27	6	1.4	21	Sequence 1, Appl

28 6 1.4 21 1 US-08-285-661-1 Sequence 1, Appl
29 6 1.4 21 1 US-08-301-838-1 Sequence 1, Appl
30 6 1.4 21 1 US-08-389-487-4 Sequence 4, Appl
31 6 1.4 21 1 US-08-389-487-12 Sequence 12, Appl
32 6 1.4 21 1 US-08-507-124-2 Sequence 2, Appl
33 6 1.4 21 1 US-08-507-124-3 Sequence 3, Appl
34 6 1.4 21 1 US-08-342-931-1 Sequence 1, Appl
35 6 1.4 21 1 US-08-400-256-1 Sequence 1, Appl
36 6 1.4 21 2 US-08-508-664-9 Sequence 9, Appl
37 6 1.4 21 2 US-08-353-476-85 Sequence 85, Appl
38 6 1.4 21 2 US-08-353-476-87 Sequence 87, Appl
39 6 1.4 21 2 US-08-484-219-1 Sequence 1, Appl
40 6 1.4 21 2 US-08-379-587-1 Sequence 1, Appl
41 6 1.4 21 2 US-08-992-676-1 Sequence 1, Appl
42 6 1.4 21 2 US-08-992-676-6 Sequence 6, Appl
43 6 1.4 21 2 US-09-134-836-1 Sequence 1, Appl
44 6 1.4 21 2 US-09-134-836-7 Sequence 7, Appl
45 6 1.4 21 4 US-08-932-082-1 Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-07-942-245-98
; Sequence 98, Application US/07942245
; Patent No. 5639641
; GENERAL INFORMATION:
; APPLICANT: PEDERSEN, Jan T.
; APPLICANT: SEARLE, Stephen M.J.
; APPLICANT: REES, Anthony R.
; APPLICANT: ROGUSKA, Michael A.
; APPLICANT: GUILD, Braydon C.
; TITLE OF INVENTION: SURFACE RESIDUE VENEERING OF RODENT
; TITLE OF INVENTION: ANTIBODIES
; NUMBER OF SEQUENCES: 522
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Hp 9000/700 Workstation
; OPERATING SYSTEM: UNIX
; SOFTWARE: In house
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/942,245
; FILING DATE: 09-SEP-1992
; CLASSIFICATION: 530
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-942-245-98

Query Match 1.6%; Score 7; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 387 PRSGKA 393
|||||

Db 4 PRSGKA 10

RESULT 2
US-09-889-14
; Sequence 14, Application US/09097889
; Patent No. 6218117
; GENERAL INFORMATION:
; APPLICANT: HerinStadt, Corrina
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Davis, Robert E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
; TITLE OF INVENTION: AGENTS THAT QUANTITATIVELY ALTER DETECTABLE
; TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA: MITOCHONDRIAL DNA RATIOS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/097,889
; FILING DATE: 15-JUN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rosenman Ph.D., Stephen J.
; REGISTRATION NUMBER: 43,058
; REFERENCE/DOCKET NUMBER: 660088.417
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; TOPOLOGY: linear
US-09-097-889-14

Query Match 1.6%; Score 7; DB 4; Length 347;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LTLISLS 14
Db 155 LTLISLS 161

RESULT 3
US-08-231-729B-6
; Sequence 6, Application US/08231729B
; Patent No. 5618722
; GENERAL INFORMATION:
; APPLICANT: ZENNO, Shuhei
; APPLICANT: SHIRAIISHI, Shinji
; APPLICANT: INOUE, Satoshi
; APPLICANT: SAIGO, Kaoru
; TITLE OF INVENTION: FIREFLY LUCIFERASE GENE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEYDIG, VOIT & MAYER
; STREET: 700 Thirteenth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,729B
; FILING DATE: 20-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 119050/1993
; FILING DATE: 21-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rose, Herbert C.
; REGISTRATION NUMBER: 29846
; REFERENCE/DOCKET NUMBER: 60130/No. 5618722aka
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-737-6770
; TELEFAX: 202-737-6776
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 552 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-231-729B-6

Query Match 1.6%; Score 7; DB 1; Length 552;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 LSFDLDR 181
Db 184 LSFDLDR 190

RESULT 4
US-08-890-865A-4
; Sequence 4, Application US/08890865A
; Patent No. 6307019
; GENERAL INFORMATION:
; APPLICANT: Constantini, Franklin
; APPLICANT: Zeng, Li
; TITLE OF INVENTION: AXIN GENE AND USES THEREOF
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/890,865A
; FILING DATE: 10-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/54249
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)278-0400
; TELEFAX: (212)391-0526
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 900 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-890-865A-4

Query Match 1.6%; Score 7; DB 4; Length 900;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 403 KGISGY 409
| | | | |
Db 268 KGISGY 274

RESULT 5

US-08-215-805A-80
; Sequence 80, Application US/08215805A
; Patent No. 5559008
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; TITLE OF INVENTION: LEUKOTOXIN GENE FROM PASTEURILLA
; NUMBER OF SEQUENCES: 84
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/215,805A
; FILING DATE: 22-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Timian, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/61 (D-1329A)
; TELEPHONE: (716) 263-1636
; TELEFAX: (716) 263-1600

; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 934 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Pasteurella suis
; STRAIN: 5943
; IMMEDIATE SOURCE:
; LIBRARY: P. suis DNA in Bacteriophage lambda-dash
; CLONE: (Lambda)yfc33-37
US-08-215-805A-80

Query Match 1.6%; Score 7; DB 1; Length 934;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 238 VKAAEDL 244
| | | | |
Db 62 VKAAEDL 68

RESULT 6

US-08-542-003-6

; Sequence 6, Application US/08542003
; Patent No. 5864013
; GENERAL INFORMATION:
; APPLICANT: Goldberg, Edward B.
; TITLE OF INVENTION: MATERIALS FOR THE PRODUCTION OF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie and Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/542,003
; FILING DATE: 13-OCT-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 8471-0005-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: 212-869-8864
; TELEX: 66441 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1026 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Bacteriophage T4
; IMMEDIATE SOURCE:
; CLONE: p37 amino acid
US-08-542-003-6

Query Match 1.6%; Score 7; DB 2; Length 1026;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 128 QTKGKP 134
| | | | |
Db 861 QTKGKP 867

RESULT 7

US-08-322-760A-6
; Sequence 6, Application US/08322760A
; Patent No. 5877279
; GENERAL INFORMATION:
; APPLICANT: Goldberg, Edward B.
; TITLE OF INVENTION: MATERIALS FOR THE PRODUCTION OF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie and Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/322,760A
FILING DATE: 13-OCT-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 8471-0003-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: 212-869-8864
TELEX: 66441 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1026 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Bacteriophage T4
IMMEDIATE SOURCE:
CLONE: p37 amino acid
US-08-322-760A-6

Query Match 1.6%; Score 7; DB 2; Length 1026;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 QTIKGP 134
|||||||
DB 861 QTIKGP 867

RESULT 8
US-08-974-775-30
Sequence 30, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
APPLICANT: Toback, F. Gary
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775
FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid

STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-30

Query Match 1.4%; Score 6; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
|||||||
DB 5 NHEATS 10

RESULT 9
US-08-974-775-29
Sequence 29, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
APPLICANT: Toback, F. Gary
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775
FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-29

Query Match 1.4%; Score 6; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
|||||||
DB 5 NHEATS 10

RESULT 10
US-08-974-775-28
Sequence 28, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:

APPLICANT: Toback, F. Gary
APPLICANT: Walsh-Reitz, Margaret
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
STREET: Plaza Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775
FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-28

Query Match 1.4%; Score 6; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 424 NHEATS 429
Db 5 NHEATS 10

RESULT 11
US-08-974-775-8
Sequence 8, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
APPLICANT: Toback, F. Gary
APPLICANT: Walsh-Reitz, Margaret
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
STREET: Plaza Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775

FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-8

Query Match 1.4%; Score 6; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 424 NHEATS 429
Db 8 NHEATS 13

RESULT 12
US-08-974-775-27
Sequence 27, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
APPLICANT: Toback, F. Gary
APPLICANT: Walsh-Reitz, Margaret
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
STREET: Plaza Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775
FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-27

Query Match 1.4%; Score 6; DB 3; Length 13;

Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

QY 424 NHEATS 429
Db 5 NHEATS 10

RESULT 13

US-08-974-775-9
; Sequence 9, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-974-775-9

Query Match 1.4%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

QY 424 NHEATS 429
Db 8 NHEATS 13

RESULT 14

US-08-974-775-10
; Sequence 10, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-974-775-9

Query Match 1.4%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

QY 424 NHEATS 429
Db 8 NHEATS 13

RESULT 15

US-08-974-775-32
; Sequence 32, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:


```

; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 14
; OTHER INFORMATION: /product= "Ser or Tyr"
US-08-974-775-32

```

```

Query Match      1.4%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 424 NHEATS 429
Db 8 NHEATS 13

```

Search completed: May 9, 2002, 12:55:41
Job time: 551 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:46:30 ; Search time 44.37 Seconds
(without alignments)
758.827 Million cell updates/sec

Title: US-09-787-083-2

Perfect score: 442

Sequence: 1 MKVSLSTLTLSILSCFAIIA.....YNHEATSFVGLMLNDWMGL 442

Scoring table:

OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

PIR_68.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	2.0	382	2 E81195	phospholipase A1,
2	9	2.0	409	2 H81831	probable phospholip
3	8	1.8	286	2 B36371	outer membrane pho
4	8	1.8	299	2 T47989	RAV-like protein -
5	8	1.8	1396	2 S36851	L-shaped tail fibe
6	7	1.6	100	2 S61050	probable membrane
7	7	1.6	122	2 E75377	hypothetical prote
8	7	1.6	158	2 E86498	Pts IIA Protein [i
9	7	1.6	158	2 E72124	pts iia protein -
10	7	1.6	197	2 A54506	tubulin alpha-1 ch
11	7	1.6	219	2 T19438	hypothetical prote
12	7	1.6	240	2 A61544	tubulin alpha chai
13	7	1.6	264	2 A44959	coat protein - pot
14	7	1.6	267	2 A60366	coat protein - pot
15	7	1.6	267	2 S14001	genome polyprotein
16	7	1.6	267	2 J26630	capsid protein - p
17	7	1.6	267	2 JC1527	coat protein - pot
18	7	1.6	267	2 S13239	tail fiber protein
19	7	1.6	267	2 E84092	phage-related prot
20	7	1.6	270	2 S73334	abc transport ATP-
21	7	1.6	273	2 C72328	transaminase B hom
22	7	1.6	287	2 T72421	oligopeptide ABC t
23	7	1.6	288	2 T36237	probable phosphotr
24	7	1.6	306	2 T27985	hypothetical prote
25	7	1.6	310	2 T02200	probable RAV-like
26	7	1.6	313	2 JT0960	polyprotein - pota
27	7	1.6	321	2 H82240	conserved hypothet
28	7	1.6	324	2 T23876	hypothetical prote
29	7	1.6	326	2 A43939	proteinase inhibit

30	7	1.6	338	2 F69437	hypothetical prote
31	7	1.6	341	2 T47653	pectate lyase-like
32	7	1.6	342	2 G64411	hypothetical prote
33	7	1.6	342	2 C64394	hypothetical prote
34	7	1.6	347	1 DNHUN2	NADH dehydrogenase
35	7	1.6	347	2 T11051	NADH dehydrogenase
36	7	1.6	347	2 B59153	NADH dehydrogenase
37	7	1.6	349	2 T43920	yfuc protein [impo
38	7	1.6	370	2 G83219	probable dihydroli
39	7	1.6	379	2 S13556	genome polyprotein
40	7	1.6	381	2 G82132	ribonuclease D vcl
41	7	1.6	382	2 S13237	tail fiber protein
42	7	1.6	384	2 S70638	tubulin alpha chai
43	7	1.6	384	2 S70641	tubulin alpha chai
44	7	1.6	384	2 S70640	tubulin alpha chai
45	7	1.6	387	2 S33517	tubulin alpha chai

ALIGNMENTS

RESULT 1

E81195

phospholipase A1, probable NMB0464 [imported] - Neisseria meningitidis (strain MC58 s
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: E81195

R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.

ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.

Science 287, 1809-1815, 2000

A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;

A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A;Reference number: A81000; MUID:20175755

A;Accession: E81195

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-382 <TET>

A;Cross-references: GB:AE002403; GB:AE002098; MID:g7225688; PIDN:AAF40501.1; PID:g722

A;Experimental source: serogroup B, strain MC58

C;Genetics:

A;Gene: NMB0464

Query Match 2.0% Score 9; DB 2; Length 382;
Best Local Similarity 100.0%; Pred. No. 0.45; Mismatches 0; Indels 0; Caps 0;
Matches 9; Conservative 0;

Qy 279 PEIFLTQPV 287

|||||||

Db 221 PEIFLTQPV 229

RESULT 2

H81831

probable phospholipase NMA2021 [imported] - Neisseria meningitidis (strain Z2491 serog
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C;Accession: H81831

R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo

; Holroyd, S.; Jørgensen, K.; Leather, S.; Mungall, K.; Quail, M.A.; Rajandre

Nature 404, 502-506, 2000

A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491

A;Reference number: A81775; MUID:20222556

A;Accession: H81831

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-409 <PAR>

A;Cross-references: GB:AL162757; GB:AL157959; MID:g7380371; PIDN:CAB85240.1; PID:g738

A;Experimental source: serogroup A, strain Z2491

C;Genetics:

A;Gene: NMA2021

Query Match
Best Local Similarity 2.0%; Score 9; DB 2; Length 409;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIFLTQPV 287
|||||
DB 248 PEIFLTQPV 256

RESULT 3
B36971
outer membrane phospholipase A (EC 3.1.1.-) precursor - Klebsiella pneumoniae
C:Species: Klebsiella pneumoniae
C>Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 18-Jun-1999
C:Accession: B36971; S40129
R:Brok, R.G.P.M.; Brinkman, E.; van Bortel, R.; Bekkers, A.C.A.P.A.; Verheij, H.M.; Tomm
J. Bacteriol. 176, 861-870, 1994
A:Title: Molecular characterization of enterobacterial pldA genes encoding outer membran
A:Reference number: A36971; MUID:94131966
A:Accession: B36971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-286 <BRO>
A:CROSS-references: EMBL:X76901; NID:g436880; PIDN:CAA54223.1; PID:g436881
A:Note: authors translated the codon AAG for residue 112 as Arg
C:Genetics:
A:Gene: pldA
C:Superfamily: bacterial phospholipase A1
C:Keywords: carboxylic ester hydrolase

Query Match
Best Local Similarity 1.8%; Score 8; DB 2; Length 286;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 431 GVGGLMLND 438
|||||
DB 277 GVGGLMLND 284

RESULT 4
T47989
RAV-like protein - Arabidopsis thaliana
N:Alternate names: protein F21F14.140
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47989
R:Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sa
submitted to the Protein Sequence Database, February 2000
A:Reference number: Z24481
A:Accession: T47989
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-299 <CHO>
A:CROSS-references: EMBL:AL138642
A:Experimental source: cultivar Columbia; BAC clone F21F14
C:Genetics:
A:Map position: 3
A:Note: F21F14.140

Query Match
Best Local Similarity 1.8%; Score 8; DB 2; Length 299;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 SLSFOLDL 181
|||||
DB 292 SLSFOLDL 299

RESULT 5
S36851

L-shaped tail fiber protein - phage T5
N:Alternate names: ltf protein
C:Species: phage T5
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 17-Nov-2000
C:Accession: S65934; S01984; S36851
R:Kallman, A.V.; Kulshin, V.E.; Shlyapnikov, M.G.; Ksenzenko, V.N.; Kryukov, V.M.
FEBS Lett. 366, 46-48, 1995
A:Title: The nucleotide sequence of the bacteriophage T5 ltf gene.
A:Reference number: S65934; MUID:95309401
A:Accession: S65934
A:Molecule type: DNA
A:Residues: 1-1396 <KAL>
A:CROSS-references: EMBL:X69460; NID:g15415; PIDN:CAA49220.1; PID:g15416
R:Kallman, A.V.; Kryukov, V.M.; Bayev, A.A.
Nucleic Acids Res. 16, 6230, 1988
A:Title: The nucleotide sequence of bacteriophage T5 DNA at the region between early
A:Reference number: S01982; MUID:88289370
A:Accession: S01984
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 934-985, 'A' 987-1396 <KA2>
A:CROSS-references: EMBL:X07559
C:Genetics:
A:Gene: ltf
C:Keywords: late protein; tail fiber

Query Match
Best Local Similarity 1.8%; Score 8; DB 2; Length 1396;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 QSATQSAS 61
|||||
DB 87 QSATQSAS 94

RESULT 6
S61050
probable membrane protein YDL163w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein D1505
C:Species: Saccharomyces cerevisiae
C>Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 20-Jun-2000
C:Accession: S61050; S67715
R:Pohl, T.M.
submitted to the EMBL Data Library, November 1995
A:Reference number: S61010
A:Accession: S61050
A:Molecule type: DNA
A:Residues: 1-100 <POH>
A:CROSS-references: EMBL:Z67750; NID:g1061256; PID:g1061276
R:Pohl, T.M.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67708
A:Accession: S67715
A:Molecule type: DNA
A:Residues: 1-100 <POW>
A:CROSS-references: EMBL:Z74212; NID:g1431258; PID:g1431260; MIPS:YDL163w
A:Experimental source: strain S288C
C:Genetics:
A:Map position: 4L
C:Superfamily: Saccharomyces probable membrane protein YDL163w
C:Keywords: transmembrane protein
F:3-19/Domain: transmembrane #status predicted <TMM>

Query Match
Best Local Similarity 1.6%; Score 7; DB 2; Length 100;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LTLSTLS 14
|||||
DB 26 LTLSTLS 32

RESULT 7

E75377
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: E75377
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: E75377
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-122 <WHI>
A:Cross-references: GB:AE002002; GB:AE000513; NID:g6459345; PIDN:AAF11150.1; PID:g6459345
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1580
A:Map position: 1

Query Match 1.6%; Score 7; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 70 EHEPELY 76
|||||
Db 115 EHEPELY 121

RESULT 8

E86498
Pts IIA Protein [imported] - Chlamydomophila pneumoniae (strain J138)
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: E86498
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349
A:Accession: E86498
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-158 <STO>
A:Cross-references: GB:BA000008; NID:g8978434; PIDN:BAA98271.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: ptsN_1
C:Superfamily: phosphotransferase system enzyme II; phosphotransferase system mannitol-s

Query Match 1.6%; Score 7; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LSTLTLS 11
|||||
Db 126 LSTLTLS 132

RESULT 9

E72124
pts iia protein - Chlamydomophila pneumoniae (strain CWL029)
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
C:Accession: E72124
R:Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606

A:Accession: E72124
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-158 <ARN>
A:Cross-references: GB:AE001591; GB:AE001363; NID:g4376311; PIDN:AAD18213.1; PID:g437
A:Experimental source: strain CWL029
C:Genetics:
A:Gene: ptsN_1
C:Superfamily: phosphotransferase system enzyme II; phosphotransferase system mannitol-specific enzyme II factor III hom
F;28-156/Domain: phosphotransferase system mannitol-specific enzyme II factor III hom

Query Match 1.6%; Score 7; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LSTLTLS 11
|||||
Db 126 LSTLTLS 132

RESULT 10

A54506
tubulin alpha-1 chain - Plasmodium yoelii (fragment)
C:Species: Plasmodium yoelii
C:Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 13-Aug-1999
C:Accession: A54506
R:Kellia, R.; Arasu, P.; Vaidya, A.B.
Mol. Biochem. Parasitol. 30, 165-174, 1988
A:Title: Molecular clones of alpha-tubulin genes of Plasmodium yoelii reveal an unusu
A:Reference number: A54506; MUID:89014607
A:Accession: A54506
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-197 <AKE>
A:Cross-references: GB:M29816; NID:g160729; PIDN:AAA29779.1; PID:g160730
C:Superfamily: tubulin

Query Match 1.6%; Score 7; DB 2; Length 197;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 114 IKTKRSI 120
|||||
Db 82 IKTKRSI 88

RESULT 11

T19438
hypothetical protein C25A1.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T19438
R:Mortimore, B.
Submitted to the EMBL Data Library, October 1996
A:Reference number: Z19124
A:Accession: T19438
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-219 <WIL>
A:Cross-references: EMBL:Z81038; PIDN:CAB02762.1; GSPDB:GN000019; CESP:C25A1.1
A:Experimental source: clone C25A1
C:Genetics:
A:Gene: CESP:C25A1.1
A:Map position: 1
A:Introns: 33/1; 66/1; 151/2

Query Match 1.6%; Score 7; DB 2; Length 219;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 158 KDAKQLE 164
|||||||
Db 52 KDAKQLE 58

RESULT 12

A61544
tubulin alpha chain - common octopus (fragment)
C:Species: Octopus vulgaris (common octopus)
C:Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 10-Jul-1998
C:Accession: A61544
R:Zinov'eva, R.D.; Aleinikova, K.S.; Tomarev, S.I.
Dokl. Akad. Nauk SSSR 302, 462-467, 1988
A:Title: Isolation and structural characterization of cDNAs coding for alpha-tubulin of
A:Reference number: A61544

A:Accession: A61544
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-240 <ZIN>
C:Superfamily: tubulin
F:239-240/Cleavage site: Glu-Tyr (tubulin-specific carboxypeptidase) #status predicted
F:239-240/Cross-link: peptide (Glu-Tyr) (by tubulin-tyrosine ligase) #status predicted

Query Match 1.6%; Score 7; DB 2; Length 240;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120
|||||||
Db 124 IKTKRSI 130

RESULT 13

A44959
coat protein - potato virus Y (strain N) (fragment)
C:Species: potato virus Y, PVY
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Nov-2000
C:Accession: A44959
R:Hay, J.M.; Fellowes, A.P.; Timmerman, G.M.
Arch. Virol. 107, 111-122, 1989
A:Title: Nucleotide sequence of the coat protein gene of a necrotic strain of potato virus
A:Reference number: A44959; MUID:90025790

A:Accession: A44959
A:Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-264 <HAY>
A:Cross-references: GB:M22470; NID:9333508; PIDN:AAA47182.1; PID:g3333509
C:Superfamily: tobacco etch virus genome polyprotein

Query Match 1.6%; Score 7; DB 2; Length 264;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 TKKDAKQ 162
|||||||
Db 8 TKKDAKQ 14

RESULT 14

A60366
coat protein - potato virus Y (strain TH)
C:Species: potato virus Y, PVY
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Nov-2000
C:Accession: A60366
R:Hataya, T.; Sano, T.; Ohshima, K.; Shikata, E.
Virus Genes 4, 339-350, 1990
A:Title: Polymerase chain-reaction-mediated cloning and expression of the coat protein gene
A:Reference number: A60366; MUID:91143125

A:Accession: A60366
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-267 <HAT>
C:Superfamily: tobacco etch virus genome polyprotein

Query Match 1.6%; Score 7; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 TKKDAKQ 162
|||||||
Db 11 TKKDAKQ 17

RESULT 15

S14001
genome polyprotein - potato virus Y (strain H) (fragment)
N:Contains: coat protein
C:Species: potato virus Y, PVY
A:Variety: strain H
C:Date: 21-Nov-1993 #sequence_revision 08-Nov-1996 #text_change 17-Nov-2000
C:Accession: S14001
R:Dalmay, T.; Balazs, E.
Nucleic Acids Res. 18, 6721, 1990
A:Title: Nucleotide sequence of an altered virulence potato virus Y coat protein gene
A:Reference number: S14001; MUID:91067494

A:Accession: S14001
A:Status: translation not shown
A:Molecule type: genomic RNA
A:Residues: 1-267 <DAL>
A:Cross-references: EMBL:X54611; NID:g61449; PIDN:CAA38432.1; PID:g930257
A:Experimental source: strain H
C:Superfamily: tobacco etch virus genome polyprotein
C:Keywords: coat protein; polyprotein
F:1-267/Product: coat protein #status predicted <CPR>

Query Match 1.6%; Score 7; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 TKKDAKQ 162
|||||||
Db 11 TKKDAKQ 17

Search completed: May 9, 2002, 12:56:32
Job time: 602 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:55:45 ; Search time 28.06 Seconds
(without alignments)
577.543 Million cell updates/sec

Title: US-09-787-083-2

Perfect score: 442

Sequence: 1 MKVSLSTLTLSCFAILA.....YNHEATSPGVGLMLNDWML 442

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 100059 seqs, 36664827 residues

Word size : 0

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	1.8	286	1	PAL_KLEPN
2	8	1.8	1396	1	VLRF_BET5
3	7	1.6	186	1	PRL_ICTPU
4	7	1.6	197	1	TBA_PLAYO
5	7	1.6	240	1	TBA_OCTVU
6	7	1.6	270	1	Y304_MYCPN
7	7	1.6	273	1	ILVE_THEMA
8	7	1.6	322	1	MIAA_PSEPU
9	7	1.6	326	1	SPI_BACBR
10	7	1.6	342	1	FLAD_METJA
11	7	1.6	342	1	Y755_METJA
12	7	1.6	347	1	NU2M_CAPHI
13	7	1.6	347	1	NU2M_HUMAN
14	7	1.6	347	1	NU2M_PANTR
15	7	1.6	347	1	NU2M_PONPA
16	7	1.6	347	1	NU2M_SHEEP
17	7	1.6	423	1	AMT_AQUAE
18	7	1.6	444	1	TBA_AUCKE
19	7	1.6	448	1	TBA4_HUMAN
20	7	1.6	448	1	TBA5_CHICK
21	7	1.6	449	1	TBA2_DROME
22	7	1.6	449	1	TBAE_PHYPO
23	7	1.6	449	1	TBAN_PHYPO
24	7	1.6	449	1	TBA_XENLA
25	7	1.6	450	1	TBAD_PHYPO
26	7	1.6	450	1	TBA_HAECO
27	7	1.6	450	1	TBA_NOTVI
28	7	1.6	451	1	TBAL_CRIGR
29	7	1.6	451	1	TBA3_HOMAM
30	7	1.6	451	1	TBA_OCTDO
31	7	1.6	451	1	TBA_TORNA
32	7	1.6	511	1	LGT_CITUN
33	7	1.6	513	1	ATPA_HAEIN

ALIGNMENTS

RESULT 1

ID	PAL_KLEPN	STANDARD;	PRT;	286 AA.
AC	P37446;			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT			
DE	PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-			
DE	ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).			
GN	PLDA.			
OS	Klebsiella pneumoniae.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Klebsiella.			
OX	NCBI_TaxID=573;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94131966; PubMed=8300539;			
RA	Brok R.G.P.M., Brinkman E., van Bortel R., Bekkers A.C.A.P.,			
RA	Verheij H.M., Tomassen J.;			
RT	"Molecular characterization of enterobacterial plda genes encoding			
RT	outer membrane phospholipase A.;"			
RL	J. Bacteriol. 176:861-870(1994).			
CC	-1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE			
CC	A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.			
CC	-1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O -> 1-ACYLGLYCERO-			
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.			
CC	-1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O -> 2-ACYLGLYCERO-			
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.			
CC	-1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.			
CC	-1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES			
CC	LOCATED THERE.			

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC	EMBL; X76901; CAA54223.1; -.
DR	PIR; B36971; B36971.
DR	PIR; S40129; S40129.
DR	InterPro; IPR003187; PLAL.
DR	Pfam; PF02253; PLAL; 1.
KW	Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium.
FT	SIGNAL 1 20 BY SIMILARITY.
FT	CHAIN 21 286 PHOSPHOLIPASE A1.
FT	ACT_SITE 161 161 BY SIMILARITY.
SQ	SEQUENCE 286 AA; 32544 MW; 3E39F863085108A3 CRC64;

Query Match 1.88; Score 8; DB 1; Length 286;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Oy 431 GVGMLND 438
      |||||
Db 277 GVGMLND 284

RESULT 2
VLTF_BPT5 STANDARD; PRT; 1396 AA.
AC P13390; Q48502;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE L-SHAPED TAIL FIBER PROTEIN (LTF PROTEIN).
GN LTF.
OS Bacteriophage T5.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae.
OC NCBI_TaxID=10726;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95309401; PubMed=7789514;
RA Kaliman A.V.; Kulshin V.E.; Shlyapnikov M.G.; Ksenzenko V.N.;
RA Kryukov V.M.;
RT "The nucleotide sequence of the bacteriophage T5 ltf gene.";
RL FEBS Lett. 366:46-48(1995).
[2]
RP SEQUENCE FROM N.A.
RA Kaliman A.V.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
[3]
RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=88289370; PubMed=3267228;
RA Kaliman A.V.; Kryukov V.M.; Bayev A.A.;
RT "The nucleotide sequence of bacteriophage T5 DNA at the region
between early and late genes.";
RL Nucleic Acids Res. 16:6230-6230(1988).
CC -1- FUNCTION: NONESSENTIAL PROTEIN THAT MEDIATES BINDING TO THE
POLYMANNOSE O ANTIGEN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X69460; CAA49220.1; -.
DR EMBL; AJ001191; CAA04591.1; -.
DR PIR; S01982; S01982.
KW Late protein.
FT CONFLICT 986 986 V -> A (IN REF. 2).
SQ SEQUENCE 1396 AA; 147989 MW; 18CD2192F65FFFC1 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 1396;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 54 QSATQSAS 61
      |||||
Db 87 QSATQSAS 94

RESULT 3
PRL_ICTPU STANDARD; PRT; 186 AA.
AC P51904;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE PROLACTIN (PRL).
OS Ictalurus punctatus (Channel catfish).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
Siluriformes; Ictaluridae; Ictalurus.
NCBI_TaxID=7998;
[1]
RP SEQUENCE.
RX TISSUE=Pituitary;
RX MEDLINE=93364578; PubMed=1308206;
RA Watanabe K.; Igarashi A.; Noso T.; Chen T.T.; Dunham R.A.;
RA Kawauchi H.;
RT "Chemical identification of catfish growth hormone and prolactin.";
RL Mol. Mar. Biol. Biotechnol. 1:239-249(1992).
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
HSSP; Q28632; 1AN3
DR InterPro; IPR001400; SOMATOTROPIN.
DR Pfam; PF00103; Hormone; 1.
DR PRINTS; PR00836; SOMATOTROPIN.
DR PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary.
FT DISULFID 45 159 BY SIMILARITY.
FT DISULFID 176 186 BY SIMILARITY.
SQ SEQUENCE 186 AA; 20606 MW; 7BCCD89718F44E74 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 186;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 41 NDLGQDN 47
      |||||
Db 140 NDLGQDN 146

RESULT 4
TBA_PLAYO STANDARD; PRT; 197 AA.
AC P12543;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE TUBULIN ALPHA CHAIN (FRAGMENT).
OS Plasmodium berghei yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OC NCBI_TaxID=5862;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89014607; PubMed=2459618;
RA Akella R.; Arasu P.; Vaidya A.B.;
RT "Molecular clones of alpha-tubulin genes of Plasmodium yoelii reveal
an unusual feature of the carboxy terminus.";
RL Mol. Biochem. Parasitol. 30:165-174(1988).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; M29816; AAA29779.1; -.
DR PIR; A54506; A54506.
DR InterPro; IPR000217; Tubulin.
DR InterPro; IPR003008; Tubulin_Ftsz.
DR Pfam; PF00091; tubulin; 1.
DR PROSITE; PS00227; TUBULIN; PARTIAL.
```

```
KW Microtubules; GTP-binding.
FT NON_TER 1
SQ SEQUENCE 197 AA; 22111 MW; 4DAAF199CC6CD319 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 197;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 IKTKRSI 120
DB 82 IKTKRSI 88

RESULT 5
TBA_OCTVU STANDARD; PRT; 240 AA.
AC P24635;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE TUBULIN ALPHA CHAIN (FRAGMENT).
OS Octopus vulgaris (Octopus).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;
OC Incirrata; Octopodidae; Octopus.
ON NCBI_TaxID=6645;
RX [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lens;
RA Zinov'Eva R.D., Aleinikova K.S., Tomarev S.I.;
RT "Isolation and structural characterization of cDNAs coding for alpha-
tubulin of the octopus eye lens.";
RL Dokl. Akad. Nauk SSSR 302:462-467(1988).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15845; CAA33844.1; -.
DR PIR; A61544; A61544.
DR InterPro; IPR000217; Tubulin.
DR InterPro; IPR003008; Tubulin_FtsZ.
DR Pfam; PF00091; tubulin; 1.
DR PROSITE; PS00227; TUBULIN; PARTIAL.
KW Microtubules; GTP-binding.
FT NON_TER 1
FT SITE 240 240 INVOLVED IN POLYMERIZATION.
SQ SEQUENCE 240 AA; 26961 MW; 13BB3A1F740F2416 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 IKTKRSI 120
DB 124 IKTKRSI 130

RESULT 6
Y304_MYCPN STANDARD; PRT; 270 AA.
ID Y304_MYCPN
AC P75355;
DT 01-NOV-1997 (Rel. 35, Created)
```

```
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOPHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN MG304 HOMOLOG
DE (A05_ORF270L).
GN MPN433 OR MP408.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000040; AAB96056.1; -.
DR InterPro; IPR003593; AAA.
DR InterPro; IPR003439; ABC_transportr.
DR InterPro; IPR001687; ATP_GTP_A.
DR Pfam; PF00005; ABC_tran; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW Hypothetical protein; ATP-binding; Transport; Complete proteome.
FT NP_BIND 36 43 ATP (POTENTIAL).
SQ SEQUENCE 270 AA; 30770 MW; 6512640B4BC051B4 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 270;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 KQFTPLS 174
DB 125 KQFTPLS 131

RESULT 7
ILVE_THEME STANDARD; PRT; 273 AA.
ID ILVE_THEME
AC P74921;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROBABLE BRANCHED-CHAIN AMINO ACID AMINOTRANSFERASE (EC 2.6.1.42)
DE (BCAT).
GN ILVE OR TM0831.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
```

RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RL genome sequence of *Thermotoga maritima*.";
 RN Nature 399:323-329(1999).
 RP [2]
 RC SEQUENCE OF 68-273 FROM N.A.
 RX STRAIN-MSB8 / DSM 3109;
 RA MEDLINE=97017137; PubMed=8863738;
 RA Guipaud O., Labedan B., Forterre P.;
 RT "A gyrB-like gene from the hyperthermophilic bacterium *Thermotoga*
 RL *maritima*.";
 RL Gene 174:121-128(1996).
 CC -1- CATALYTIC ACTIVITY: L-LEUCINE + 2-OXOGlutARATE = 4-METHYL-2-
 CC OXOPENTANOATE + L-GLUTAMATE (ALSO ACTS ON L-ISOLEUCINE AND
 CC L-VALINE).
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -1- PATHWAY: VALINE AND ISOLEUCINE BIOSYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO CLASS-IV OF PYRIDOXAL-PHOSPHATE-DEPENDENT
 CC AMINOTRANSFERASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE001750; AAC35913.1; -
 DR EMBL; AU9692; AAC44497.1; -
 DR HSPSP; P00510; IA3G.
 DR TIGR; TM0831; -
 DR InterPro; IPR001544; AminoTran_4.
 DR Pfam; PF01063; aminotran_4; 1.
 DR ProDom; PD001961; AminoTran_4; 1.
 DR PROSITE; PS00770; AA_TRANSF_CLASS.4; 1.
 DR Transfaser; AminoTransferase; Branched-chain amino acid biosynthesis;
 KW Pyridoxal phosphate; Complete proteome.
 FT CONFLICT 77 A -> R (IN REF. 2).
 SQ SEQUENCE 273 AA; 31158 MW; 2162B705612E90E3 CRC64;

 Query Match 1.6%; Score 7; DB 1; Length 273;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 172 PLSLSD 178
 Db 61 PLSLSD 67

 RESULT 8
 ID MIAA_PSEPU STANDARD; PRT; 322 AA.
 AC 030762;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE TRNA DELTA(2)-ISOPENTENYLPROPHOSPHATE TRANSFERASE (EC 2.5.1.8) (IPP
 DE TRANSFERASE) (ISOPENTENYL-DIPHOSPHATE:TRNA ISOPENTENYLTTRANSFERASE)
 DE (IPTASE) (IPT).
 GN MIAA.
 OS Pseudomonas putida.
 CC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 CC Pseudomonas.
 CC NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M;
 RA Olekhnovich I.N., Gussin G.N.;
 RA "Attenuation of the *Pseudomonas putida* trpE and trpGDC genes.";
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CATALYZES THE FIRST STEP IN THE BIOSYNTHESIS OF
 CC 2-METHYLTHIO-N6-(DELTA(2)-ISOPENTENYL)-ADENOSINE (MS[2][6]A)

CC ADJACENT TO THE ANTICODON OF SEVERAL TRNA SPECIES (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ISOPENTENYL DIPHOSPHATE + TRNA -
 CC PYROPHOSPHATE + TRNA CONTAINING 6-ISOPENTENYLADENOSINE.
 CC -1- SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF016312; AAB69443.1; -
 DR InterPro; IPR002627; IPPT.
 DR Pfam; PF01715; IPPT; 1.
 DR ProDom; PD004674; IPPT; 1.
 KW Transfaser; Nucleotidyltransferase; trna processing; ATP-binding.
 FT NP_BIND 12 19 ATP (POTENTIAL).
 SQ SEQUENCE 322 AA; 35488 MW; DIAA81A2D7B32F6A CRC64;

 Query Match 1.6%; Score 7; DB 1; Length 322;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 33 FVDEVRS 39
 Db 231 FVDEVRS 237

 RESULT 9
 ID SPI_BACBR STANDARD; PRT; 326 AA.
 AC P43131;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE PROTEASE INHIBITOR PRECURSOR (BBRPI).
 DE Bacillus brevis.
 OS Bacteria; Firmicutes; Bacillus/Clostridium group;
 CC Bacillus/Staphylococcus group; Brevibacillus.
 CC NCBI_TaxID=1393;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-32; 104-112 AND 122-136.
 RC STRAIN=HPD31;
 RX MEDLINE=92304060; PubMed=1610177;
 RA Shiga Y., Hasegawa K., Tsuboi A., Yamagata H., Uda S.;
 RT "Characterization of an extracellular protease inhibitor of *Bacillus*
 RL *brevis* HPD31 and nucleotide sequence of the corresponding gene.";
 RL Appl. Environ. Microbiol. 58:525-531(1992).
 CC -1- FUNCTION: SHOWS INHIBITORY ACTIVITY TOWARDS SERINE PROTEASES, SUCH
 CC AS TRYPSIN, CHYMOTRYPSIN, AND SUBTILISIN. MAY FORM A TRYPSIN-
 CC INHIBITOR COMPLEX IN A MOLAR RATIO OF 1:1. IT IS HEAT RESISTANT AT
 CC NEUTRAL AND ACIDIC PHs.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- PTM: MAY UNDERGO ACTIVATION AFTER SECRETION. IT IS PRODUCED
 CC EXTRACELLULARLY IN MULTIPLE FORMS HAVING AT LEAST THREE DIFFERENT
 CC MOLECULAR WEIGHTS (BBRPI-A, -B, AND -C).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D10696; BAA01538.1; -
 KW Serine protease inhibitor; Protease inhibitor; Signal; Repeat.
 FT SIGNAL 1 24
 FT CHAIN 25 326 SERINE PROTEASE INHIBITOR.
 FT CHAIN 104 326 SERINE PROTEASE INHIBITOR-C.

```

FT CHAIN 104 ? SERINE PROTEASE INHIBITOR-B (POTENTIAL).
FT CHAIN 122 326 SERINE PROTEASE INHIBITOR-A.
FT DOMAIN 177 304 CONTAINS TWO APPROXIMATE REPEATS.
FT REPEAT 177 208 1.
FT REPEAT 272 304 2.
SQ SEQUENCE 326 AA; 35100 MW; 1C0456ABFA912E77 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 326;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 PVAFVDE 36
Db 119 PVAFVDE 125

RESULT 10
FLAD_METJA STANDARD; PRT; 342 AA.
AC Q58305;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PUTATIVE FLAGELLA-RELATED PROTEIN D.
GN FLAD OR MJ0895.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67521; AAB98748.1; -
DR TIGR; MJ0755; -
KW Hypothetical protein; Signal; Complete proteome.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 342 HYPOTHETICAL PROTEIN MJ0755.
SQ SEQUENCE 342 AA; 38778 MW; BD25220A7EEAD85C CRC64;

Query Match 1.6%; Score 7; DB 1; Length 342;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 YTTALEN 82
Db 90 YTTALEN 96

RESULT 12
NU2M_CAPHI STANDARD; PRT; 347 AA.
ID NU2M_CAPHI
AC Q36346;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 2 (EC 1.6.5.3).
GN MTND2 OR ND2.
OS Capra hircus (Goat).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RC DOVC P., Mann W., Hecht W.;
RT Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.
RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE -> NAD(+) + UBIQUINOL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN; MITOCHONDRIAL

```

CC INNER MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X72965; CAA51468.1; -
DR InterPro; IPR003917; NADHUB_oxdrdctse2.
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; oxidored_q1; 1
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
SQ SEQUENCE 347 AA; 39136 MW; 6129CFF90980475E CRC64;

Query Match 1.6%; Score 7; DB 1; Length 347;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 LTLTSLIS 14
Db 155 LTLTSLIS 161

RESULT 13
NU2M_HUMAN STANDARD; PRT; 347 AA.
AC P03891; Q34769; Q9TGI0; Q9TGI2; Q9TGI3; Q9TGI4;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 2 (EC 1.6.5.3).
GN MTND2 OR ND2.
OS Homo sapiens (Human).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81173052; PubMed=7219534;
RA Anderson S., Bankier A.T., Barrell B.G., de Bruijn M.H.L.,
RA Coulson A.R., Drouin J., Eperon I.C., Nierlich D.P., Roe B.A.,
RA Sanger F., Schreier P.H., Smith A.J.H., Staden R., Young I.G.;
RT "Sequence and organization of the human mitochondrial genome.";
RL Nature 290:457-465(1981).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=81170577; PubMed=6260957;
RA Sanger F., Coulson A.R., Barrell B.G., Smith A.J.H., Roe B.A.;
RT "Cloning in single-stranded bacteriophage as an aid to rapid DNA
RT sequencing.";
RL J. Mol. Biol. 143:161-178(1980).
RN [3]
RP SEQUENCE FROM N.A. AND VARIANTS.
RX MEDLINE=98133898; PubMed=9475751;
RA Wise C.A., Stram M., Eastel S.;
RT "Departure from neutrality at the mitochondrial NADH dehydrogenase
RT subunit 2 gene in humans, but not in chimpanzees.";
RL Genetics 148:409-421(1998).
RN [4]
RP SEQUENCE FROM N.A. AND VARIANTS ILE-43; LEU-325 AND THR-331.
RC TISSUE=Placenta;
RA MEDLINE=95132634; PubMed=7530363;
RA Horai S., Hayasaka K., Kondo R., Tsugane K., Takahata N.;
RT "Recent African origin of modern humans revealed by complete sequences
RT of hominoid mitochondrial DNAs";
RL Proc. Natl. Acad. Sci. U.S.A. 92:532-536(1995).
RN [5]

RP IDENTIFICATION OF PROTEIN.
RX MEDLINE=85188293; PubMed=3921850;
RA Chomyn A., Mariottini P., Cleeter M.W.J., Ragan C.I., Matsuno-Yagi A.,
RA Hatefi Y., Doolittle R.F., Attardi G.;
RT "Six unidentified reading frames of human mitochondrial DNA encode
RT components of the respiratory-chain NADH dehydrogenase.";
RL Nature 314:592-597(1985).
RN [6]
RP VARIANT LHON ASP-150.
RX MEDLINE=91144615; PubMed=1900003;
RA Johns D.R., Berman J.;
RT "Alternative, simultaneous complex I mitochondrial DNA mutations in
RT Leber's hereditary optic neuropathy";
RL Biochem. Biophys. Res. Commun. 174:1324-1330(1991).
RN [7]
RP VARIANTS L-42; R-63; A-119; P-148; S-150; T-159 AND A-185.
RX MEDLINE=92098084; PubMed=1757091;
RA Marzuki S., Noer A.S., Lertrit P., Thyagarajan D., Kapsa R.,
RA Utthanaphol P., Byrne E.;
RT "Normal variants of human mitochondrial DNA and translation products:
RT the building of a reference data base";
RL Hum. Genet. 88:139-145(1991).
RN [8]
RP VARIANT LHON SER-259.
RX MEDLINE=92120513; PubMed=1732158;
RA Brown M.D., Voljavec A.S., Lott M.T., Torroni A., Yang C.C.,
RA Wallace D.C.;
RT "Mitochondrial DNA complex I and III mutations associated with
RT Leber's hereditary optic neuropathy";
RL Genetics 130:163-173(1992).
RN [9]
RP VARIANT AD SER-331.
RX MEDLINE=92118019; PubMed=1370613;
RA Lin F.-H., Lin R., Wisniewski H.M., Hwang Y.-W., Grundke-Iqbal I.,
RA Healy-Louie G., Iqbal K.;
RT "Detection of point mutations in codon 331 of mitochondrial NADH
RT dehydrogenase subunit 2 in Alzheimer's brains";
RL Biochem. Biophys. Res. Commun. 182:238-246(1992).
RN [10]
RP VARIANT THR-57.
RX MEDLINE=98127994; PubMed=9461455;
RA Rieder M.J., Taylor S.L., Tobe V.O., Nickerson D.A.;
RT "Automating the identification of DNA variations using quality-based
RT fluorescence re-sequencing: analysis of the human mitochondrial
RT genome";
RL Nucleic Acids Res. 26:967-973(1998).
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE -> NAD(+) + UBIQUINOL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE.
CC -1- DISEASE: DEFECTS IN MTND2 ARE ONE OF THE CAUSES OF LEBER'S
CC HEREDITARY OPTIC NEUROPATHY (LHON), A MATERNALLY INHERITED DISEASE
CC RESULTING IN ACUTE BILATERAL BLINDNESS DUE TO RETINAL DEGENERATION
CC PREDOMINANTLY IN YOUNG MEN. CARDIAC CONDUCTION DEFECTS AND
CC NEUROLOGICAL DEFECTS HAVE ALSO BEEN DESCRIBED. RESULTING IN OPTIC
CC NERVE DEGENERATION AND CARDIAC DYSRHYTHMIA.
CC -1- DISEASE: DEFECTS IN MTND2 COULD BE ASSOCIATED WITH ALZHEIMER'S
CC DISEASE (AD).
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J01415; AAB58944.1; -
DR EMBL; M05462; CAA24027.1; -
DR EMBL; M10546; AAA65502.1; ALT_INIT.
DR EMBL; D38112; BAA07291.1; -
DR EMBL; AF014882; AAC25441.1; -
DR EMBL; AF014884; AAC25443.1; -

```
DR EMBL; AF014885; AAC25444.1; -
DR EMBL; AF014887; AAC25446.1; -
DR EMBL; AF014889; AAC25448.1; -
DR EMBL; AF014890; AAC25449.1; -
DR EMBL; AF014891; AAC25450.1; -
DR EMBL; AF014892; AAC25451.1; -
DR EMBL; AF014895; AAC25454.1; -
DR EMBL; AF014896; AAC25455.2; -
DR EMBL; AF014897; AAC25456.1; -
DR EMBL; AF014898; AAC25457.1; -
DR EMBL; AF014899; AAC25458.2; -
DR EMBL; AF014900; AAC25459.1; -
DR EMBL; AF014901; AAC25460.1; -
DR PIR; A00414; DNHUN2.
DR MIM; 502500; -
DR MIM; 516001; -
DR MIM; 535000; -
DR InterPro; IPR003917; NADHub_oxdrdtset2.
DR Pfam; IPR001750; Oxidored_q1.
DR Pfam; PF00361; oxidored_q1; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane;
KW Disease mutation; Leber's hereditary optic neuropathy;
KW Alzheimer's disease; Polymorphism.
FT VARIANT 42 42 P -> L.
FT VARIANT 43 43 /FTId-VAR_008590.
FT VARIANT 57 57 /FTId-VAR_011348.
FT VARIANT 63 63 I -> T.
FT VARIANT 63 63 Q -> R.
FT VARIANT 69 69 /FTId-VAR_008592.
FT VARIANT 88 88 I -> V.
FT VARIANT 88 88 N -> S.
FT VARIANT 119 119 /FTId-VAR_011350.
FT VARIANT 119 119 T -> A.
FT VARIANT 148 148 /FTId-VAR_008593.
FT VARIANT 148 148 S -> P.
FT VARIANT 150 150 /FTId-VAR_008594.
FT VARIANT 150 150 N -> D (IN LHON; SECONDARY MUTATION; DOES
FT NOT SEEM TO DIRECTLY CAUSE THE DISEASE).
FT VARIANT 150 150 /FTId-VAR_004755.
FT VARIANT 159 150 N -> S.
FT VARIANT 159 159 /FTId-VAR_008595.
FT VARIANT 185 185 I -> T.
FT VARIANT 185 185 /FTId-VAR_008596.
FT VARIANT 237 237 T -> A.
FT VARIANT 237 237 /FTId-VAR_008597.
FT VARIANT 259 259 L -> M.
FT VARIANT 259 259 G -> S (IN LHON; RARE PRIMARY MUTATION).
FT VARIANT 265 265 A -> T.
FT VARIANT 265 265 /FTId-VAR_011352.
FT VARIANT 265 265 A -> V.
FT VARIANT 278 278 /FTId-VAR_011353.
FT VARIANT 325 325 I -> T.
FT VARIANT 325 325 F -> L.
FT VARIANT 331 331 /FTId-VAR_011355.
FT VARIANT 331 331 A -> S (IN AD).
FT VARIANT 331 331 /FTId-VAR_004758.
FT VARIANT 333 333 A -> T.
FT VARIANT 333 333 /FTId-VAR_004757.
FT SEQUENCE 347 AA; 38961 MW; C06FD982317C3F2D CRC64;
Query Match 1.6%; Score 7; DB 1; Length 347;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 8 LTLISLS 14
Db 155 LTLISLS 161
RESULT 14
ID NU2M_PANTR STANDARD; PRT; 347 AA.
AC O21798;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 2 (EC 1.6.5.3).
GN MTND2 OR ND2.
OS Pan troglodytes (Chimpanzee).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP MEDLINE=98133898; PubMed=9475751;
RA Wise C.A., Stram M., Eastel S.;
RT "Departure from neutrality at the mitochondrial NADH dehydrogenase
RL Genetics 148:409-421(1998).
CC -I- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE.
CC -I- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF014908; AAC25467.1; -
DR EMBL; AF014909; AAC25468.1; -
DR EMBL; AF014910; AAC25469.1; -
DR EMBL; AF014911; AAC25470.1; -
DR EMBL; AF014912; AAC25471.1; -
DR EMBL; AF014913; AAC25472.1; -
DR EMBL; AF014914; AAC25473.1; -
DR EMBL; AF014915; AAC25474.1; -
DR EMBL; AF014916; AAC25475.1; -
DR EMBL; AF014917; AAC25476.1; -
DR EMBL; AF014918; AAC25477.1; -
DR EMBL; AF014919; AAC25478.1; -
DR EMBL; AF014920; AAC25479.1; -
DR EMBL; AF014921; AAC25480.1; -
DR InterPro; IPR003917; NADHub_oxdrdtset2.
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; oxidored_q1; 1.
KW Oxidoreductase; NAD; ubiquinone; Mitochondrion; Transmembrane.
FT VARIANT 94 94 P -> S (IN STRAIN A-292).
FT SEQUENCE 347 AA; 39020 MW; 2E8269D105810D4E CRC64;
Query Match 1.6%; Score 7; DB 1; Length 347;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 8 LTLISLS 14
Db 155 LTLISLS 161
```

```
RESULT 15
NU2M_PONPA STANDARD; PRT; 347 AA.
ID NU2M_PONPA
```

AC P92691;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 2 (EC 1.6.5.3).
 GN MTND2 OR ND2 OR NADH2.
 OS Pongo pygmaeus abelii (Sumatran orangutan).
 OC Mitochondrion.
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
 OX NCBI_TaxID=9601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YN93-312;
 RX MEDLINE=97032590; PubMed=8875856;
 RA Xu X., Arnason U.;
 RT "The mitochondrial DNA molecule of Sumatran orangutan and a molecular
 RT proposal for two (Bornean and Sumatran) species of orangutan.";
 RL J. Mol. Evol. 43:431-437(1996).
 CC -|- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
 CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 CC INNER MEMBRANE.
 CC -|- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 2 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X97707; CAA66284.1; -;
 DR InterPro: IPR003917; NADHub_oxdrdctse2.
 DR InterPro: IPR001750; Oxidored_q1.
 DR Pfam: PF00361; oxidored_q1; 1.
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
 SQ SEQUENCE 347 AA; 38619 MW; 53885266466A2B93 CRC64;

Query Match 1.68; Score 7; DB 1; Length 347;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 LTLSILS 14
 |||||
 Db 155 LTLSILS 161

Search completed: May 9, 2002, 13:05:51
 Job time: 606 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:55:05 ; Search time 77.39 Seconds
(without alignments)
835.410 Million cell updates/sec

Title: US-09-787-083-2
Perfect score: 442
Sequence: 1 MKVSLSTLTLSCFAILA.....YNHEATSGVGLMDWML 442

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 473505 seqs, 146272329 residues

Word size : 0

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- SPTREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	2.0	382	2 Q9K0U7	Q9K0U7 neisseria m
2	9	2.0	409	2 Q9J721	Q9J721 neisseria m
3	9	2.0	862	2 Q9CL62	Q9CL62 pasteurella
4	9	2.0	1432	10 Q9AX85	Q9AX85 oryza sativ
5	8	1.8	299	10 Q9W268	Q9W268 arabidopsis
6	8	1.8	306	2 Q9CL22	Q9CL22 pasteurella
7	7	1.6	55	12 Q67608	Q67608 tomato gold
8	7	1.6	72	12 Q9IZS5	Q9IZS5 potato viru
9	7	1.6	72	12 Q9IZS4	Q9IZS4 potato viru
10	7	1.6	95	12 Q9IZS3	Q9IZS3 potato viru
11	7	1.6	96	12 Q9IZS8	Q9IZS8 potato viru
12	7	1.6	100	3 Q12148	Q12148 saccharomyc
13	7	1.6	101	12 Q9IZS6	Q9IZS6 potato viru
14	7	1.6	122	2 Q9RU16	Q9RU16 deinococcus
15	7	1.6	124	12 Q9IZS7	Q9IZS7 potato viru
16	7	1.6	131	6 Q9GMD4	Q9GMD4 macaca fusc
17	7	1.6	138	12 Q9IZQ7	Q9IZQ7 potato viru
18	7	1.6	147	5 Q36035	Q36035 hexamita in
19	7	1.6	158	2 Q9Z9C1	Q9Z9C1 chlamydia p

20	7	1.6	169	2	Q9L2A7	Q9L2A7 streptomyce
21	7	1.6	186	10	Q9LXR9	Q9LXR9 arabidopsis
22	7	1.6	189	13	Q9DFT2	Q9DFT2 notothenia
23	7	1.6	198	4	Q9BUX9	Q9BUX9 homo sapien
24	7	1.6	212	13	Q9I819	Q9I819 ictalurus p
25	7	1.6	214	5	Q9VTR5	Q9VTR5 drosophila
26	7	1.6	216	2	Q9X6H6	Q9X6H6 streptococc
27	7	1.6	218	2	Q9RF14	Q9RF14 streptococc
28	7	1.6	219	5	O17594	O17594 caenorhabdi
29	7	1.6	220	2	O9XCK1	O9XCK1 streptococc
30	7	1.6	220	5	O36036	O36036 spironucleu
31	7	1.6	221	12	Q85002	Q85002 alifalra mos
32	7	1.6	227	5	Q9BMF4	Q9BMF4 glossina mo
33	7	1.6	233	6	Q9WZB0	Q9WZB0 ovis aries
34	7	1.6	236	13	Q9DFS8	Q9DFS8 notothenia
35	7	1.6	264	12	Q85260	Q85260 potato viru
36	7	1.6	265	9	Q38155	Q38155 bacterioph
37	7	1.6	267	2	Q9K732	Q9K732 bacillus ha
38	7	1.6	267	9	Q99363	Q99363 bacterioph
39	7	1.6	267	12	Q85261	Q85261 potato viru
40	7	1.6	267	12	Q98628	Q98628 potato viru
41	7	1.6	267	12	Q86842	Q86842 potato viru
42	7	1.6	267	12	Q99026	Q99026 potato viru
43	7	1.6	267	12	Q05476	Q05476 potato viru
44	7	1.6	267	12	O09711	O09711 potato viru
45	7	1.6	267	12	Q9QNM9	Q9QNM9 potato viru

ALIGNMENTS

RESULT 1
Q9K0U7
ID Q9K0U7 PRELIMINARY; PRT; 382 AA.
AC Q9K0U7;
DT 01-OCT-2000 (TREMREL. 15, Created)
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)
DT 01-JUN-2001 (TREMREL. 17, Last annotation update)
DE PHOSPHOLIPASE A1, PUTATIVE.
GN NMB0464;
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citterone H., Clark E.B.,
RA Cotton M.D., Uterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlatti V., Masignani V., Pizzo M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58";
RL Science 287:1809-1815(2000).
DR EMBL: AE002403; AAF40901.1; -.
DR TIGR: NMB0464; -.
DR InterPro: IPR003187; PLAI.
DR Pfam: PF02253; PLAI; 1.
KW Complete proteome.SQ SEQUENCE 382 AA; 42714 MW; B468A802F062E836 CRC64;

Query Match 2.0%; Score 9; DB 2; Length 382;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 279 PEIFLTQPV 287

DB 221 PEIFLTQPV 229

```

RESULT 2
Q9J721 ID Q9J721 PRELIMINARY; PRT; 409 AA.
AC Q9J721;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE PUTATIVE PHOPHOLIPASE.
GN NMA2021.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis 22491."
RL Nature 404:502-506(2000).
DR EMBL: AL162757; CAB85240.1; -.
DR InterPro; IPR003187; PLAL.
DR Pfam; PF02253; PLAL; 1.
KW Complete proteome.
SQ SEQUENCE 409 AA; 45862 MW; CD6585B064D01A41 CRC64;

Query Match 2.0%; Score 9; DB 2; Length 409;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIFLTQPV 287
Db 248 PEIFLTQPV 256

RESULT 3
Q9CL62 ID Q9CL62 PRELIMINARY; PRT; 862 AA.
AC Q9CL62;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CVGSY.
GN CVGSY OR PM1380.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RA "Complete genomic sequence of Pasteurella multocida Pm70."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -I- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.
CC -I- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC KINASES.
DR EMBL: AE006176; AAK03464.1; -.
DR InterPro; IPR000410; Bctrl_sensor.
DR InterPro; IPR000658; DUF5.
DR InterPro; IPR003660; HAMP.
DR InterPro; IPR003594; HATPase_c.
DR InterPro; IPR003661; His_kinA.

InterPro; IPR001789; Response_reg.
DR Pfam; PF00672; DUF5; 1.
DR Pfam; PF00072; response_reg; 1.
DR Pfam; PF00512; signal; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; HSKA; 1.
DR SMART; SM00448; REC; 1.
KW Complete proteome; Kinase; Phosphorylation; Sensory transduction;
KW Transferase.
SQ SEQUENCE 862 AA; 97846 MW; AB6814A3B63626C0 CRC64;

Query Match 2.0%; Score 9; DB 2; Length 862;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LSTLTLSIL 13
Db 293 LSTLTLSIL 301

RESULT 4
Q9AX85 ID Q9AX85 PRELIMINARY; PRT; 1432 AA.
AC Q9AX85;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE PUTATIVE ABC TRANSPORTER PROTEIN.
GN P0410E03.7.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0410E03."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002844; BAB21276.1; -.
SQ SEQUENCE 1432 AA; 160234 MW; 04E8CF2CAA08D993 CRC64;

Query Match 2.0%; Score 9; DB 10; Length 1432;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 FTPLSLSF 178
Db 830 FTPLSLSF 838

RESULT 5
Q9M268 ID Q9M268 PRELIMINARY; PRT; 299 AA.
AC Q9M268;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE RAV-LIKE PROTEIN.
GN F21F14.140.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
```

RA Choisine N., Robert C., Brottier P., Wincker P., Wincker P., Cattolico L.,
RA Artiguenave F., Saurin W., Weissenbach J., Mewes H.W., Lemcke K.,
RA Mayer K.F.X., Quetier F., Salanoubat M.,
RN Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; ALI38642; CAB71904.1; -
DR InterPro; IPR003340; B3.
DR Pfam; PF02362; B3; 1.
SQ SEQUENCE 299 AA; 34270 MW; 2AB841F179DD174B CRC64;

Query Match 1.8%; Score 8; DB 10; Length 299;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 174 SLSFDLDR 181
|
Db 292 SLSFDLDR 299
|

RESULT 6
Q9CLZ2 PRELIMINARY; PRT; 306 AA.
ID Q9CLZ2
AC Q9CLZ2
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL PROTEIN PM1426.
GN PM1426.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; AE006179; AAK03510.1; -
DR InterPro; IPR003187; PLAI.
DR Pfam; PF02253; PLAI; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 306 AA; 35580 MW; EAF3DE8C1C22B26E CRC64;

Query Match 1.8%; Score 8; DB 2; Length 306;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 314 KLSRSWNR 321
|
Db 184 KLSRSWNR 191
|

RESULT 7
Q67608 PRELIMINARY; PRT; 55 AA.
ID Q67608
AC Q67608
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SUBGENOMIC DNA DERIVED FROM DNA B CCDS - COVALENTLY CLOSED CIRCULAR
DE DOUBLE-STRANDED MOLECULE.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10631;
RN [1]
RP SEQUENCE FROM N.A.

RA MEDLINE=87040767; PubMed=3022243;
RA MacDowell S.W., Coutts R.H.A., Buck K.W.;
RT "Molecular characterisation of subgenomic single-stranded and double-
RT stranded DNA forms isolated from plants infected with tomato golden
RT mosaic virus.";
RL Nucleic Acids Res. 14:7967-7984(1986).
DR EMBL; X04485; CAA28171.1; -
DR InterPro; IPR00211; Gemini_BL.
DR Pfam; PF00845; Gemini_BL1; 1.
SQ SEQUENCE 55 AA; 6281 MW; 41963570D739EC96 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 55;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 312 SAKLSRS 318
|
Db 43 SAKLSRS 49
|

RESULT 8
Q9IZS5 PRELIMINARY; PRT; 72 AA.
ID Q9IZS5
AC Q9IZS5
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY(NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=122280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUBER NECROSING (NTN);
RA Cerovska N., Filigarova M., Moravec T., Petrzik K.;
RT "Differences in nucleotide and amino acid sequences of N-terminal
RT parts of coat proteins among isolates of potato virus Y-NTN strain.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF28633; AAF67874.1; -
FT NON_TER 1
FT NON_TER 72
SQ SEQUENCE 72 AA; 7721 MW; 2C5287E15624CBDD CRC64;

Query Match 1.6%; Score 7; DB 12; Length 72;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 156 TKKDAKQ 162
|
Db 11 TKKDAKQ 17
|

RESULT 9
Q9IZS4 PRELIMINARY; PRT; 72 AA.
ID Q9IZS4
AC Q9IZS4
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY(NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=122280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUBER NECROSING (NTN);
RA Cerovska N., Filigarova M., Moravec T., Petrzik K.;

RT "Differences in nucleotide and amino acid sequences of N-terminal
RT parts of coat proteins among isolates of potato virus Y-NTN strain.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF228634; AAF67875.1; -.
FT NON_TER 1 1
SQ SEQUENCE 72 AA; 7692 MW; 4F129B870523B482 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 72;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKQ 162
Db 11 TKKDAKQ 17
|||||

RESULT 10
Q91ZS3 PRELIMINARY; PRT; 95 AA.
AC Q91ZS3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY(NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=122280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TUBER NECROSING (NTN);
RA Cerovska N., Filligatova M., Moravec T., Petrzik K.;
RT "Differences in nucleotide and amino acid sequences of N-terminal
RT parts of coat proteins among isolates of potato virus Y-NTN strain.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF228635; AAF67876.1; -.
FT NON_TER 1 1
SQ SEQUENCE 95 AA; 10446 MW; 28C06891777C9928 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 95;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKQ 162
Db 11 TKKDAKQ 17
|||||

RESULT 11
Q91ZS8 PRELIMINARY; PRT; 96 AA.
AC Q91ZS8;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y (strain N) (PVY).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=12219;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NECROSING (N);
RA Cerovska N., Filligatova M., Moravec T., Petrzik K.;
RT "Differences in nucleotide and amino acid sequences of N-terminal
RT parts of coat proteins among isolates of potato virus Y-NTN strain.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF228630; AAF67871.1; -.
FT NON_TER 1 1
SQ SEQUENCE 96 AA; 10559 MW; 89C3A5675B9784CA CRC64;

Query Match 1.6%; Score 7; DB 12; Length 96;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKQ 162
Db 11 TKKDAKQ 17
|||||

RESULT 12
Q12148 PRELIMINARY; PRT; 100 AA.
AC Q12148;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE ORF YDL163W.
GN YDL163W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Pohl T.M.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MIPS;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-ALPHAS288C;
RA Pohl T.M.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: 274212; CAA98738.1; -.
DR EMBL: 267750; CAA91583.1; -.
DR SGD; S0002322; YDL163W.
KW Hypothetical protein.
SQ SEQUENCE 100 AA; 10751 MW; 952FB92C852AA710 CRC64;

Query Match 1.6%; Score 7; DB 3; Length 100;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LTLSTLS 14
Db 26 LTLSTLS 32
|||||

RESULT 13
Q91ZS6 PRELIMINARY; PRT; 101 AA.
AC Q91ZS6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY(NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=122280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TUBER NECROSING (NTN);

RA Cerovska N., Filigarova M., Moravec T., Petrzik K.;
RT "Differences in nucleotide and amino acid sequences of N-terminal
RL parts of coat proteins among isolates of potato virus Y-NTN strain.";
DR EMBL: AF228632; AAF67873.1; -
FT NON_TER 101 101
FT NON_TER 101 101
SQ SEQUENCE 101 AA; 11112 MW; B55BF51C3D3FE9BD CRC64;

Query Match 1.68; Score 7; DB 12; Length 101;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKDKAQ 162
Db 11 TKDKAQ 17
|||||
RESULT 14
Q9RU16 PRELIMINARY; PRT; 122 AA.
AC Q9RU16
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL 13.7 KDA PROTEIN.
GN DR1580.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=RI;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1."
RL Science 286:1571-1577(1999).
DR EMBL: AE002002; AAF11150.1; -
DR TIGR: DR1580;
DR InterPro: IPR000325; Glyoxalase_1.
DR DR Pfam: PF00903; Glyoxalase; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 122 AA; 13683 MW; 7BA836DCAD786E51 CRC64;

Query Match 1.68; Score 7; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 EHEPELY 76
Db 115 EHEPELY 121
|||||
RESULT 15
Q9IZS7 PRELIMINARY; PRT; 124 AA.
AC Q9IZS7
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY(NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;

OC Potyvirus.
OX NCBI_TaxID=122280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUBER NECROSING (NTN);
RA Cerovska N., Filigarova M., Moravec T., Petrzik K.;
RT "Differences in nucleotide and amino acid sequences of N-terminal
RL parts of coat proteins among isolates of potato virus Y-NTN strain.";
DR EMBL: AF228631; AAF67872.1; -
FT NON_TER 124 124
FT NON_TER 124 124
SQ SEQUENCE 124 AA; 13654 MW; DE3618D620BA7080 CRC64;

Query Match 1.68; Score 7; DB 12; Length 124;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKDKAQ 162
Db 11 TKDKAQ 17
|||||
Search completed: May 9, 2002, 13:05:17
Job time: 612 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:54:58 ; Search time 72.68 Seconds
(without alignments)
450.473 Million cell updates/sec

Title: US-09-787-083-4
Perfect score: 442
Sequence: 1 MKVSTLTLSTILPCFAILA.....YNHEATFGVGLMLNDWMGL 442

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0
Searched: 522463 seqs, 74073290 residues
Word size : 0
Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_1101.*
1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	442	100.0	442	21	AAV85269 BASB034 amino acid
2	390	88.2	442	21	AAV85270 BASB034 amino acid
3	289	65.4	442	21	AAV85271 BASB034 amino acid
4	224	50.7	442	21	AAV85268 BASB034 amino acid
5	9	2.0	370	21	AAV75156 Neisseria gonorrhoe
6	9	2.0	370	21	AAV75157 Neisseria meningit
7	9	2.0	370	21	AAV75158 Neisseria meningit
8	9	2.0	374	21	AAV70629 Neisseria meningit
9	9	2.0	375	21	AAV70628 Neisseria meningit
10	8	1.8	278	21	AAV11137 Arabidopsis thalia
11	8	1.8	283	21	AAV11136 Arabidopsis thalia

12	1.8	299	21	AAV11135 Arabidopsis thalia
13	1.6	16	15	AAV52127 Mouse light chain
14	1.6	50	21	AAV79483 Rat alpha tubulin
15	1.6	64	21	AAV06322 Arabidopsis thalia
16	1.6	82	21	AAV01488 Human secreted pro
17	1.6	106	21	AAV07608 Arabidopsis thalia
18	1.6	106	21	AAV51952 Arabidopsis thalia
19	1.6	108	21	AAV37811 Arabidopsis thalia
20	1.6	116	21	AAV87313 Human signal pepti
21	1.6	136	21	AAV37810 Arabidopsis thalia
22	1.6	169	20	AAV34661 Chlamydia pneumoni
23	1.6	181	21	AAV37809 Arabidopsis thalia
24	1.6	186	21	AAV07607 Arabidopsis thalia
25	1.6	186	21	AAV51951 Arabidopsis thalia
26	1.6	188	21	AAV51950 Arabidopsis thalia
27	1.6	190	21	AAV07606 Arabidopsis thalia
28	1.6	268	13	AAV21542 Envelope protein o
29	1.6	302	13	AAV24254 Protease inhibitor
30	1.6	316	21	AAV39500 Arabidopsis thalia
31	1.6	324	21	AAV39499 Arabidopsis thalia
32	1.6	326	13	AAV24356 Protease inhibitor
33	1.6	365	21	AAV39498 Arabidopsis thalia
34	1.6	381	20	AAV60241 Human endometrium
35	1.6	448	21	AAV91960 Human cytoskeleton
36	1.6	523	21	AAV44285 Arabidopsis thalia
37	1.6	552	16	AAV67007 Firefly luciferase
38	1.6	665	20	AAV21871 Amino acid sequenc
39	1.6	683	20	AAV21869 Amino acid sequenc
40	1.6	684	20	AAV21873 Amino acid sequenc
41	1.6	684	20	AAV21875 Amino acid sequenc
42	1.6	715	20	AAV21860 Amino acid sequenc
43	1.6	783	20	AAV70586 Human SLAP-130. H
44	1.6	900	20	AAV96264 Human axin. Homo
45	1.6	934	17	AAV07637 P. suis leukotoxin

ALIGNMENTS

RESULT 1

AAV85269
ID AAV85269 standard; Protein; 442 AA.

XX AAV85269;

XX AC AAV85269;

DT 29-JUN-2000 (first entry)

DE BASB034 amino acid sequence #2.

XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;
vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
sinusitis; nosocomial infection; invasive disease; chronic otitis media;
hearing loss; antibacterial drug.

OS Moraxella catarrhalis.

XX WO200015802-A1.

PN 23-MAR-2000.

PD 14-SEP-1999; 99WO-EP06781.

XX 14-SEP-1998; 98GB-0020002.

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

PI Ruelle J;

DR WPI; 2000-271440/23.

DR N-PSDB: AAA10701.

XX Novel BASB034 polynucleotides and polypeptides from Moraxella
catarrhalis used to prepare vaccines against bacterial infections

XX PS Claim 3; Page 67; 106pp; English.

XX CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from

CC strain Mc2908. The invention relates to BASB034 polypeptides from

CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034

CC polynucleotides and polypeptides may be employed as research reagents and

CC material for the discovery of treatments and diagnostics for diseases,

CC particularly human diseases. They are particularly used to diagnose and

CC treat M. catarrhalis infections. They can be used for diagnosis of

CC disease, staging of disease, or determining response of an infectious

CC organism to drugs. The polynucleotides may be used as a source for

CC hybridization probes, and for screening of genetic mutations, serotype,

CC organism or strain identification, identification of mutations in BASB034

CC sequences, and as components of arrays which are useful for diagnostic

CC and prognostic purposes. The polypeptides can be used to produce

CC antibodies. The polypeptides can also be used in vaccine formulations,

CC and to identify agonists and antagonists. The polypeptides, antibodies,

CC agonists and antagonists (which are bacteriostatic) are used for the

CC treatment and prevention of diseases such as otitis media in infants and

CC children, pneumonia in elderlies, sinusitis, nosocomial infections and

CC invasive diseases, and chronic otitis media with hearing loss. The

CC polypeptides, agonists and antagonists are also used for screening of

CC antibacterial drugs. The BASB034 products of the invention can be used

CC screen for new antibacterial compounds that may target resistant

CC bacteria.

XX CC Sequence 442 AA;

Query Match 100.0%; Score 442; DB 21; Length 442;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSLTLILPFAILLIOQAQVNPVAFVDEVRKNDLQDNELLIGVQATQSA 60

Db 1 mkvsltlilpfaillioqaqavnpvafvdevrskndlgdqnelligvqatqsa 60

QY 61 SYDTANPLDEHEPELYTTALENKTMNLINCSALNQDITLVHGTTPAVITKRSI 120

Db 61 stdtanpldehepeelyttalenktmlncsalnqditlvhgttpavitkrsi 120

QY 121 RDETTIWTIKGKPVVYQETDPIFLMGNEKGMITKKDAKOLEYAAKQFTPLSLSFDD 180

Db 121 ridetiwtikgkpvyvqetdtpiflmgnekgmtkkdakqleyaakqftplsfdld 180

QY 181 RNTPLWSSRRPHNPMYVLPFIEMHGKPNRSPNTPSHEARQFTNEFRAPLKFQVSVKVA 240

Db 181 rntplwssrrphnmpyvlpifimhgkpnrsnptpshearqftnefrapelkfqvsvkva 240

QY 241 AEDLMGTSDLMFGYTTQSHWQIFNGKNSRPRVRVHDYQPEIFLTQPVYSDLPWDGKVRMI 300

Db 241 aedlmgtdsdlmfgyttqshwqifngknsrprvrhdyqpeifltqpvysdlpwdgkvrmi 300

QY 301 GNGAVHHSNGESAKLSRSNNRAYLMAGMEKWNLTVMNPRIWGRIFKEGSGSQPDNDPILD 360

Db 301 gngavhhsngesaklsrsnraylmagmewknlvtmnpriwgrifkegsgsqpdndpild 360

QY 361 XYGCGVRLYOLENKNLSGVRVNRPSRGKALQLDVYVPLGKIGSGYFQIFQGYGQSL 420

Db 361 xygcvgrlyolenknlsgrvnrpsrgkalqldvvyvplgkigsgyfqi fgygqsl 420

QY 421 IDYNHEATSFVGLMLNDWMGL 442

Db 421 idynheatsfvglmlndwmgl 442

RESULT 2

ID AAY85270

XX AAY85270 standard; Protein; 442 AA.

AC AAY85270;

XX

DT 29-JUN-2000 (first entry)

XX BASB034 amino acid sequence #3.

XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;

KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;

KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;

KW hearing loss; antibacterial drug.

XX OS Moraxella catarrhalis.

XX WO200015802-A1.

XX 23-MAR-2000.

XX 14-SEP-1999; 99WO-EP06781.

XX 14-SEP-1998; 98GB-0020002.

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX Ruelle J;

XX WPI; 2000-271440/23.

DR N-PSDB; AAA10702.

XX Novel BASB034 polynucleotides and polypeptides from Moraxella

PT catarrhalis used to prepare vaccines against bacterial infections

XX Claim 3; Page 68; 106pp; English.

XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from

CC strain Mc2913. The invention relates to BASB034 polypeptides from

CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034

CC polynucleotides and polypeptides may be employed as research reagents and

CC material for the discovery of treatments and diagnostics for diseases,

CC particularly human diseases. They are particularly used to diagnose and

CC treat M. catarrhalis infections. They can be used for diagnosis of

CC disease, staging of disease, or determining response of an infectious

CC organism to drugs. The polynucleotides may be used as a source for

CC hybridization probes, and for screening of genetic mutations, serotype,

CC organism or strain identification, identification of mutations in BASB034

CC sequences, and as components of arrays which are useful for diagnostic

CC and prognostic purposes. The polypeptides can be used to produce

CC antibodies. The polypeptides can also be used in vaccine formulations,

CC and to identify agonists and antagonists. The polypeptides, antibodies,

CC agonists and antagonists (which are bacteriostatic) are used for the

CC treatment and prevention of diseases such as otitis media in infants and

CC children, pneumonia in elderlies, sinusitis, nosocomial infections and

CC invasive diseases, and chronic otitis media with hearing loss. The

CC polypeptides, agonists and antagonists are also used for screening of

CC antibacterial drugs. The BASB034 products of the invention can be used

CC screen for new antibacterial compounds that may target resistant

CC bacteria.

XX Sequence 442 AA;

Query Match 88.2%; Score 390; DB 21; Length 442;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 VQSQTASASTDTANPLDEHEPELYTTALENKTMNLINCSALNQDITLVHGTTPA 112

Db 53 vqsqtasastdtanpldehepeelyttalenktmlncsalnqditlvhgetpa 112

QY 113 VIKTKRSIRLDETIWQTIKGPQVYQETDPIFLMGNEKGMITKKDAKQLEYAAKQFTP 172

Db 113 viktkrsirldetiwtikgkpvyvqetdtpiflmgnekgmtkkdakqleyaakqfcp 172

QY 173 LSLFSDLDRNNTPLWSSRRPHNPMYVLPFIEMHGKPNRSPNTPSHEARQFTNEFRAPLKF 232

Db 173 lslfslldrnrntplwssrrphnmpyvlpifimhgkpnrsnptpshearqftnefrapelkf 232

Qy	233	QVSVKVKAAEDLWGHTDSDLWFGYTTQDSHWQIFNGKNSRFRVHDYQPELTPQVYSDLP	292
Db	233	qvsvkvkaaedlwgtdsdlwfgytqgshwqifngknsrpfvrhdyqpei fltqpyvsdlp	292
Qy	293	WDGKVRMTGMGAVHHSNGESAKLSRWNRAYLMAGWENKLNTPRIWGRIFKEGSGSQP	352
Db	293	wdgkvrmlgmavhhsngesaklsrwnraylmagmewknlcvmpriwgrifkegsgsqp	352
Qy	353	DDNPDLIDYGYGDRVFLYQLENKSNISGTVRYNPRSGKALQLDLYVPLGKGISGYFOI	412
Db	353	ddnpdlidygygdrvflqylenknsigtvrynprrsgkalgldyvpplgkgisgyfqi	412
Qy	413	FGYGSQSLIDYNHEATSFVGGLMLNDWMGL	442
Db	413	fgygsgslidynheatsfvgglmlndwmgl	442
RESULT	3		
AAy85271		ID AAY85271 standard; Protein; 442 AA.	
XX	XX	AC AAY85271;	
XX	XX		
DT	29-JUN-2000	(first entry)	
XX	XX		
DE	BASB034	amino acid sequence #4.	
XX	XX		
KW	Moraxella catarrhalis infection; BASB034; diagnosis; staging;		
KW	vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;		
KW	sinusitis; nosocomial infection; invasive disease; chronic otitis media;		
KW	hearing loss; antibacterial drug.		
XX	XX		
OS	Moraxella catarrhalis.		
XX	XX		
PN	WO2000015802-A1.		
XX	XX		
PD	23-MAR-2000.		
XX	XX		
PF	14-SEP-1999; 99WO-EP06781.		
XX	XX		
PR	14-SEP-1998; 98GB-0020002.		
XX	XX		
PA	(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.		
PI	Ruelle J;		
XX	XX		
DR	WPI; 2000-271440/23.		
DR	N-PSDB; AAA10703.		
XX	XX		
PT	Novel BASB034 polynucleotides and polypeptides from Moraxella		
PT	catarrhalis used to prepare vaccines against bacterial infections		
XX	XX		
PS	Claim 3; Page 69; 106pp: English.		
XX	XX		
CC	This sequence represents a Moraxella catarrhalis BASB034 polypeptide from		
CC	strain Mc2969. The invention relates to BASB034 polypeptides from		
CC	M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034		
CC	polynucleotides and polypeptides may be employed as research reagents and		
CC	material for the discovery of treatments and diagnostics for diseases,		
CC	particularly human diseases. They are particularly used to diagnose and		
CC	treat M. catarrhalis infections. They can be used for diagnosis of		
CC	disease, staging of disease, or determining response of an infectious		
CC	organism to drugs. The polynucleotides may be used as a source for		
CC	hybridization probes, and for screening of genetic mutations, serotype,		
CC	CC organism or strain identification, identification of mutations in BASB034		
CC	sequences, and as components of arrays which are useful for diagnostic		
CC	and prognostic purposes. The polypeptides can be used to produce		
CC	antibodies. The polypeptides can also be used in vaccine formulations,		
CC	and to identify agonists and antagonists. The polypeptides, antibodies,		
CC	agonists and antagonists (which are bacteriostatic) are used for the		
CC	treatment and prevention of diseases such as otitis media in infants and		
CC	children, pneumonia in elderlies, sinusitis, nosocomial infections and		

XX Novel BASB034 polynucleotides and polypeptides from Moraxella
PT catarrhalis used to prepare vaccines against bacterial infections
XX
PS
XX Claim 3; Fig 2; 106pp; English.
CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2931 (ATCC 43617). The invention relates to BASB034 polypeptides
CC from M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The
CC BASB034 polynucleotides and polypeptides may be employed as research
CC reagents and material for the discovery of treatments and diagnostics for
CC diseases, particularly human diseases. They are particularly used to
CC diagnose and treat M. catarrhalis infections. They can be used for
CC diagnosis of disease, staging of disease, or determining response of an
CC infectious organism to drugs. The polynucleotides may be used as a source
CC for hybridization probes, and for screening of genetic mutations,
CC serotype, organism or strain identification, identification of mutations
CC in BASB034 sequences, and as components of arrays which are useful for
CC produce antibodies. The polypeptides can also be used in vaccine
CC formulations, and to identify agonists and antagonists. The polypeptides,
CC antibodies, agonists and antagonists (which are bacteriostatic) are used
CC for the treatment and prevention of diseases such as otitis media in
CC infants and children, pneumonia in elderly, sinusitis, nosocomial
CC infections and invasive diseases, and chronic otitis media with hearing
CC loss. The polypeptides, agonists and antagonists are also used for
CC screening of antibacterial drugs. The BASB034 products of the invention
CC can be used screen for new antibacterial compounds that may target
CC resistant bacteria.
XX
SQ Sequence 442 AA;

Query Match 50.7%; Score 224; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 1.4e-209;
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 219 QFTNPEFRAPELKFQSVKVAEDLWGTDSLWFGYQQSHWQFNGKNSRPFVRHVDYQ 278
DB 219 qftnpefrapelkfqsvkvkaaedlwtgdsdlwfgyqqshwqfngknsrpfvrhdyq 278
QY 279 PEIFLTQPVYSLPDWDGVRHMGAVHSHNGESAKLSRSNRAYLMAGMEWKNLTVMR 338
DB 279 peifltqpvysldpwdgvrhmgavhshngesaklsrsnraylmagmewknlvtmvr 338
QY 339 IWRGIFKEGSGSQPDNDPILDYGYGVDRFLYQLENKSNISGTVRYNPRSGKGLQLDY 398
DB 339 lwrgifkegsgsqpdndpildygygvdvrflyqlenksnlsrgctvrynprrsgkglqldy 398
QY 399 VYPLGKGISGYFQIFQGYGQSLIDYNHEATSGVGLMLNDWMGL 442
DB 399 vyplgkgisgyfqiifqgygqslidynheatsfgvglmndwmgl 442

RESULT 5
AAV75156
ID AAV75156 standard; Protein; 370 AA.
XX
AC AAV75156;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria gonorrhoeae ORF 582 protein sequence SEQ ID NO:1786.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy.
XX
OS Neisseria gonorrhoeae.
XX
PN WO9957280-A2.
XX
PD 11-NOV-1999.

XX 30-APR-1999; 99WO-US09346.
PF
XX 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizzo M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
DR WPI: 2000-062150/05.
DR N-PSDB; AAZ53918.
XX
XX Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics -
XX
PS Claim 2; Page 903; 1453pp; English.
CC
CC AAZ53015 TO AAZ54536, AAZ54577 TO AAZ54615, and AAZ74253 TO AAZ75941
CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
CC and polypeptides. AAZ54537 TO AAZ54576 and AAZ54616 TO AAZ5473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 370 AA;
Query Match 2.0%; Score 9; DB 21; Length 370;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 279 PEIFLTQPV 287
DB 209 peifltqpv 217
RESULT 6
AAV75157
ID AAV75157 standard; Protein; 370 AA.
XX
AC AAV75157;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 582 protein sequence SEQ ID NO:1788.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy.
XX
OS Neisseria meningitidis.
XX
PN WO9957280-A2.
XX
PD 11-NOV-1999.

PF 30-APR-1999; 99WO-US09346.
 XX 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 98US-0121528.
 XX (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX

PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX WPI; 2000-062150/05.
 DR N-PSDB; AA253919.
 XX

XX Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics -
 PT

XX Claim 2; Page 904; 1453pp; English.

XX AA253015 to AA254536, AA254577 to AA254615, and AA274253 to AA275941
 CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA255473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC *Neisseria* bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of *Neisseria* bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 XX

SQ Sequence 370 AA;

Query Match 2.0%; Score 9; DB 21; Length 370;
 Best Local Similarity 100.0%; Pred. No. 2.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIFLTQPV 287
 |||||
 Db 209 peifltqpv 217

RESULT 7
 AA275158
 ID AA275158 standard; Protein; 370 AA.

XX AA275158;
 AC

XX 21-MAR-2000 (first entry)
 DT

XX *Neisseria meningitidis* ORF 582 protein sequence SEQ ID NO:1790.
 DE

XX *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
 KW antibacterial; gene therapy.
 XX

XX *Neisseria meningitidis*.
 OS

XX WO9957280-A2.
 PN

XX 11-NOV-1999.
 PD

XX 30-APR-1999; 99WO-US09346.
 PF

XX 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 98US-0121528.
 XX (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX

PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX WPI; 2000-062150/05.
 DR N-PSDB; AA253920.
 XX

XX Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics -
 PT

XX Claim 2; Page 905; 1453pp; English.

XX AA253015 to AA254536, AA254577 to AA254615, and AA274253 to AA275941
 CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA255473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC *Neisseria* bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of *Neisseria* bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 XX

SQ Sequence 370 AA;

Query Match 2.0%; Score 9; DB 21; Length 370;
 Best Local Similarity 100.0%; Pred. No. 2.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIFLTQPV 287
 |||||
 Db 209 peifltqpv 217

RESULT 8
 AA270629
 ID AA270629 standard; Protein; 374 AA.

XX AA270629;
 AC

XX 18-JUL-2000 (first entry)
 DT

XX *Neisseria meningitidis* serogroup B strain H44/76 BASB033 protein.
 DE

XX BASB033; diagnosis, prophylaxis; treatment; antibacterial; vaccine;
 KW *Neisseria meningitidis* infection.
 KW

XX *Neisseria meningitidis*.
 OS

XX WO200015801-A1.
 PN

XX 23-MAR-2000.
 PD

XX 09-SEP-1999; 99WO-EF06718.
 PF

XX 14-SEP-1998; 98GB-0020003.
 PR

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX Ruelle J;
XX WPI; 2000-271439/23.
XX N-PSDB; AAZ52134.
XX Isolated BASB033 polypeptides and polynucleotides of Neisseria
PT meningitidis, useful for diagnosis, prophylaxis and treatment of N.
PT meningitidis infection -
XX Claim 4; Page 59; 93pp; English.
XX The present sequence is a BASB033 protein from
CC Neisseria meningitidis serogroup B strain H44776. The protein
CC shows homology to the Klebsiella pneumoniae outer membrane
CC phospholipase A. The present sequence is useful for diagnosis,
CC prophylaxis and treatment of N. meningitidis infection. It may also be
CC used for the discovery and development of antibacterial compounds and
CC in vaccine compositions.
XX Sequence 374 AA;
SQ

Query Match 2.0%; Score 9; DB 21; Length 374;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 279 PEIFLTQPV 287
Db 213 pelfltqpv 221
IIIIIIII

RESULT 9
AAY70628
ID AAY70628 standard; Protein; 375 AA.
XX AC AAY70628;
XX DT 18-JUL-2000 (first entry)
XX DE Neisseria meningitidis serogroup B strain ATCC13090 BASB033 protein.
XX BASB033; diagnosis, prophylaxis; treatment; antibacterial; vaccine;
KW Neisseria meningitidis infection.
KW Neisseria meningitidis.
OS WO200015801-A1.
PN 23-MAR-2000.
PD 09-SEP-1999; 99WO-EP06718.
PF 14-SEP-1998; 98GB-0020003.
PR (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX Ruelle J;
XX WPI; 2000-271439/23.
XX N-PSDB; AAZ52133.
XX Isolated BASB033 polypeptides and polynucleotides of Neisseria
PT meningitidis, useful for diagnosis, prophylaxis and treatment of N.
PT meningitidis infection -
XX Claim 4; Page 58; 93pp; English.
XX The present sequence is a BASB033 protein from
CC Neisseria meningitidis serogroup B strain ATCC13090. The protein
CC shows homology to the Klebsiella pneumoniae outer membrane

CC phospholipase A. The present sequence is useful for diagnosis,
CC prophylaxis and treatment of N. meningitidis infection. It may also be
CC used for the discovery and development of antibacterial compounds and
CC in vaccine compositions.
XX Sequence 375 AA;
SQ

Query Match 2.0%; Score 9; DB 21; Length 375;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 279 PEIFLTQPV 287
Db 214 pelfltqpv 222
IIIIIIII

RESULT 10
AAG11137
ID AAG11137 standard; Protein; 278 AA.
XX AC AAG11137;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 9741.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
PD 06-SEP-2000.
PF 25-FEB-2000; 2000EP-0301439.
PR 25-FEB-1999; 99US-0121825.
PR 09-MAR-1999; 99US-0123180.
PR 23-MAR-1999; 99US-0123548.
PR 25-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 01-APR-1999; 99US-0126785.
PR 06-APR-1999; 99US-0127462.
PR 16-APR-1999; 99US-0128234.
PR 19-APR-1999; 99US-0128714.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 07-MAY-1999; 99US-0132486.
PR 11-MAY-1999; 99US-0132487.
PR 14-MAY-1999; 99US-0132488.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 99US-0136392.

PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140655.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142830.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0136782.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 8; DB 21; Length 278;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 SLSFOLDLR 181
 |||||||
Db 271 sisfldlr 278

RESULT 11

AA011136
ID AAG11136 standard; Protein; 283 AA.

XX
AC AAG11136;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 9740.

XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

XX
PN EP1033405-A2.

XX
PD 06-SEP-2000.

XX
PF 25-FEB-2000; 2000EP-0301439.

XX
PR 25-FEB-1999; 99US-0121825.

PR
PR 05-MAR-1999; 99US-0123180.

PR
PR 09-MAR-1999; 99US-0123548.

PR
PR 23-MAR-1999; 99US-0125788.

PR
PR 25-MAR-1999; 99US-0126264.

PR
PR 29-MAR-1999; 99US-0126785.

PR
PR 01-APR-1999; 99US-0127462.

PR
PR 06-APR-1999; 99US-0128234.

PR
PR 08-APR-1999; 99US-0128714.

PR
PR 16-APR-1999; 99US-0129845.

PR
PR 19-APR-1999; 99US-0130077.

PR
PR 21-APR-1999; 99US-0130449.

PR
PR 23-APR-1999; 99US-0130510.

PR
PR 28-APR-1999; 99US-0130891.

PR
PR 30-APR-1999; 99US-0131449.

PR
PR 30-APR-1999; 99US-0132048.

PR
PR 04-MAY-1999; 99US-0132407.

PR
PR 05-MAY-1999; 99US-0132484.

PR
PR 06-MAY-1999; 99US-0132485.

PR
PR 06-MAY-1999; 99US-0132486.

PR
PR 07-MAY-1999; 99US-0132487.

PR
PR 11-MAY-1999; 99US-0132863.

PR
PR 14-MAY-1999; 99US-0134256.

PR
PR 14-MAY-1999; 99US-0134218.

PR
PR 14-MAY-1999; 99US-0134219.

PR
PR 14-MAY-1999; 99US-0134221.

PR
PR 14-MAY-1999; 99US-0134370.

PR
PR 18-MAY-1999; 99US-0134768.

PR
PR 19-MAY-1999; 99US-0134941.

PR
PR 20-MAY-1999; 99US-0135124.

PR
PR 21-MAY-1999; 99US-0135353.

PR
PR 24-MAY-1999; 99US-0135629.

PR
PR 25-MAY-1999; 99US-0136021.

PR
PR 27-MAY-1999; 99US-0136392.

PR
PR 28-MAY-1999; 99US-0136782.

PR
PR 01-JUN-1999; 99US-0137222.

PR
PR 03-JUN-1999; 99US-0137528.

PR
PR 04-JUN-1999; 99US-0137528.

PR
PR 07-JUN-1999; 99US-0137724.

PR
PR 08-JUN-1999; 99US-0138094.

PR
PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 1-88; Score 8; DB 21; Length 283;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 174 SLSFDLDR 181
|||||||

Db 276 slsfdldr 283
RESULT 12
AAG11135
ID AAG11135 standard; Protein; 299 AA.
XX
AC AAG11135;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 9739.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147202.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158023.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 8; DB 21; Length 299;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 SLSFDLDR 181
|||||||
DB. 292 slsfdldr 299

RESULT 13
AAR52127
ID AAR52127 standard; Peptide; 16 AA.
XX

KW termination sequence.
XX Arabidopsis thaliana.
PN EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 09-MAR-1999; 99US-0123180.
XX 23-MAR-1999; 99US-0123548.
XX 25-MAR-1999; 99US-0125788.
XX 29-MAR-1999; 99US-0126264.
XX 01-APR-1999; 99US-0126785.
XX 06-APR-1999; 99US-0127462.
XX 08-APR-1999; 99US-0128234.
XX 16-APR-1999; 99US-0128714.
XX 19-APR-1999; 99US-0129845.
XX 21-APR-1999; 99US-0130077.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 04-MAY-1999; 99US-0132048.
XX 05-MAY-1999; 99US-0132484.
XX 06-MAY-1999; 99US-0132485.
XX 07-MAY-1999; 99US-0132486.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 18-MAY-1999; 99US-0134370.
XX 19-MAY-1999; 99US-0134768.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135124.
XX 24-MAY-1999; 99US-0135353.
XX 25-MAY-1999; 99US-0135623.
XX 27-MAY-1999; 99US-0136021.
XX 28-MAY-1999; 99US-0136392.
XX 01-JUN-1999; 99US-0136782.
XX 03-JUN-1999; 99US-0137222.
XX 04-JUN-1999; 99US-0137528.
XX 07-JUN-1999; 99US-0137502.
XX 08-JUN-1999; 99US-0137724.
XX 10-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 14-JUN-1999; 99US-0138847.
XX 16-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 17-JUN-1999; 99US-0139453.
XX 18-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 24-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140693.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.

PR	10-SEP-1999;	99US-0153070;
PR	13-SEP-1999;	99US-0153758;
PR	15-SEP-1999;	99US-0154018;
PR	16-SEP-1999;	99US-0154039;
PR	20-SEP-1999;	99US-0154779;
PR	22-SEP-1999;	99US-0155139;
PR	23-SEP-1999;	99US-0155486;
PR	24-SEP-1999;	99US-0155659;
PR	28-SEP-1999;	99US-0156458;
PR	29-SEP-1999;	99US-0156596;
PR	04-OCT-1999;	99US-0157117;
PR	05-OCT-1999;	99US-0157753;
PR	06-OCT-1999;	99US-0157865;
PR	07-OCT-1999;	99US-0158029;
PR	08-OCT-1999;	99US-0158232;
PR	12-OCT-1999;	99US-0158369;
PR	13-OCT-1999;	99US-0159293;
PR	13-OCT-1999;	99US-0159294;
PR	13-OCT-1999;	99US-0159295;
PR	14-OCT-1999;	99US-0159329;
PR	14-OCT-1999;	99US-0159330;
PR	14-OCT-1999;	99US-0159331;
PR	14-OCT-1999;	99US-0159637;
PR	14-OCT-1999;	99US-0159638;
PR	18-OCT-1999;	99US-0159584;
PR	21-OCT-1999;	99US-0160741;
PR	21-OCT-1999;	99US-0160767;
PR	21-OCT-1999;	99US-0160768;
PR	21-OCT-1999;	99US-0160770;
PR	21-OCT-1999;	99US-0160814;
PR	21-OCT-1999;	99US-0160815;
PR	22-OCT-1999;	99US-0160980;
PR	22-OCT-1999;	99US-0160981;
PR	25-OCT-1999;	99US-0160989;
PR	25-OCT-1999;	99US-0161404;
PR	25-OCT-1999;	99US-0161405;
PR	25-OCT-1999;	99US-0161406;
PR	26-OCT-1999;	99US-0161359;
PR	26-OCT-1999;	99US-0161360;
PR	26-OCT-1999;	99US-0161361;
PR	28-OCT-1999;	99US-0161920;
PR	28-OCT-1999;	99US-0161992;
PR	28-OCT-1999;	99US-0161993;
PR	29-OCT-1999;	99US-0162142;

Search completed: May 9, 2002, 12:54:58
Job time: 513 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:55:41 ; Search time 36.87 seconds
(without alignments)
269.771 Million cell updates/sec

Title: US-09-787-083-4

Perfect score: 442

Sequence: 1 MKVSLTLTLILPCFAILAAAA.....YNHEATSFVGLMLNDWMGL 442

Scoring table: OLIGO

Gapex 60.0 , Gapext 60.0

Searched: 212252 seqs, 22503292 residues

Word size : 0

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
1	7	1.6	16	Sequence 98, Appl
2	7	1.6	552	Sequence 6, Appl
3	7	1.6	900	Sequence 4, Appl
4	7	1.6	934	Sequence 80, Appl
5	7	1.6	1026	Sequence 6, Appl
6	7	1.6	1026	Sequence 6, Appl
7	6	1.4	10	Sequence 30, Appl
8	6	1.4	11	Sequence 28, Appl
9	6	1.4	12	Sequence 8, Appl
10	6	1.4	13	Sequence 27, Appl
11	6	1.4	14	Sequence 9, Appl
12	6	1.4	14	Sequence 10, Appl
13	6	1.4	14	Sequence 32, Appl
14	6	1.4	15	Sequence 5, Appl
15	6	1.4	16	Sequence 4, Appl
16	6	1.4	20	Sequence 1, Appl
17	6	1.4	21	Sequence 13, Appl
18	6	1.4	21	Sequence 15, Appl
19	6	1.4	21	Sequence 1, Appl
20	6	1.4	21	Sequence 1, Appl
21	6	1.4	21	Sequence 13, Appl
22	6	1.4	21	Sequence 15, Appl
23	6	1.4	21	Sequence 1, Appl
24	6	1.4	21	Sequence 1, Appl
25	6	1.4	21	Sequence 1, Appl
26	6	1.4	21	Sequence 1, Appl
27	6	1.4	21	Sequence 1, Appl

28	6	1.4	21	1	US-08-301-838-1	Sequence 1, Appl
29	6	1.4	21	1	US-08-389-487-4	Sequence 4, Appl
30	6	1.4	21	1	US-08-389-487-12	Sequence 12, Appl
31	6	1.4	21	1	US-08-507-124-2	Sequence 2, Appl
32	6	1.4	21	1	US-08-507-124-3	Sequence 3, Appl
33	6	1.4	21	1	US-08-342-931-1	Sequence 1, Appl
34	6	1.4	21	1	US-08-400-256-1	Sequence 1, Appl
35	6	1.4	21	2	US-08-508-664-9	Sequence 9, Appl
36	6	1.4	21	2	US-08-353-476-85	Sequence 85, Appl
37	6	1.4	21	2	US-08-353-476-87	Sequence 87, Appl
38	6	1.4	21	2	US-08-484-219-1	Sequence 1, Appl
39	6	1.4	21	2	US-08-979-587-1	Sequence 1, Appl
40	6	1.4	21	2	US-08-992-676-1	Sequence 1, Appl
41	6	1.4	21	2	US-08-992-676-6	Sequence 6, Appl
42	6	1.4	21	2	US-09-134-836-1	Sequence 1, Appl
43	6	1.4	21	2	US-09-134-836-7	Sequence 7, Appl
44	6	1.4	21	3	US-08-967-867-2	Sequence 2, Appl
45	6	1.4	21	4	US-08-932-082-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-07-942-245-98
; Sequence 98, Application US/07942245
; Patent No. 5639641
; GENERAL INFORMATION:
; APPLICANT: PEDERSEN, Jan T.
; APPLICANT: SEARLE, Stephen M.J.
; APPLICANT: REES, Anthony R.
; APPLICANT: ROGUSKA, Michael A.
; APPLICANT: GUILD, Braydon C.
; TITLE OF INVENTION: SURFACE RESIDUE VENEERING OF RODENT
; TITLE OF INVENTION: ANTIBODIES
; NUMBER OF SEQUENCES: 522
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughree, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: UNIX
; SOFTWARE: In house
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/942,245
; FILING DATE: 09-SEP-1992
; CLASSIFICATION: 530
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-942-245-98

Query Match 1.6%; Score 7; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 387 PRSGKA 393
Db 4 PRSGKA 10

RESULT 2
US-08-231-729B-6
; Sequence 6, Application US/08231729B
; Patent No. 5618722
; GENERAL INFORMATION:
; APPLICANT: ZENNO, Shuhei
; APPLICANT: SHIRAIISHI, Shinji
; APPLICANT: INOUE, Satoshi
; APPLICANT: SAIGO, Kaoru
; TITLE OF INVENTION: FIREFLY LUCIFERASE GENE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEYDIG, VOIT & MAYER
; STREET: 700 Thirteenth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,729B
; FILING DATE: 20-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 119050/1993
; FILING DATE: 21-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rose, Herbert C.
; REGISTRATION NUMBER: 29846
; REFERENCE/DOCKET NUMBER: 60130/No. 5618722aka
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-737-6770
; TELEFAX: 202-737-6776
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 552 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-231-729B-6

Query Match 1.6%; Score 7; DB 1; Length 552;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 175 LSFDLDR 181
Db 184 LSFDLDR 190

RESULT 3
US-08-890-865A-4
; Sequence 4, Application US/08890865A
; Patent No. 6307019
; GENERAL INFORMATION:
; APPLICANT: Constantini, Franklin
; APPLICANT: Zeng, Li
; TITLE OF INVENTION: AXIN GENE AND USES THEREOF
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10036
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/890,865A
; FILING DATE: 10-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/54249
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)278-0400
; TELEFAX: (212)391-0526
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 900 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-890-865A-4

Query Match 1.6%; Score 7; DB 4; Length 900;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 403 GKGISGY 409
Db 268 GKGISGY 274

RESULT 4
US-08-215-805A-80
; Sequence 80, Application US/08215805A
; Patent No. 5559008
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; TITLE OF INVENTION: LEUKOTOXIN GENE FROM PASTURELLA
; NUMBER OF SEQUENCES: 84
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/215,805A
; FILING DATE: 22-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Timian, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/61 (D-1329A)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1636
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 934 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```
; ORIGINAL SOURCE:
; ORGANISM: Pasteurella suis
; STRAIN: 5943
; IMMEDIATE SOURCE:
; LIBRARY: P. suis DNA in Bacteriophage lambda-dash
; CLONE: (Lambda)yfc33-37
; US-08-215-805A-80

Query Match          1.6%; Score 7; DB 1; Length 934;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 238 VKAEDL 244
Db 62 VKAEDL 68

RESULT 5
US-08-542-003-6
; Sequence 6, Application US/08542003
; Patent No. 5864013
; GENERAL INFORMATION:
; APPLICANT: Goldberg, Edward B.
; TITLE OF INVENTION: MATERIALS FOR THE PRODUCTION OF
; TITLE OF INVENTION: NANOMETER STRUCTURES AND USE THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie and Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/542,003
; FILING DATE: 13-OCT-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 8471-0005-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: 212-869-8864
; TELEX: 66441 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1026 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Bacteriophage T4
; IMMEDIATE SOURCE:
; CLONE: p37 amino acid
; US-08-542-003-6

Query Match          1.6%; Score 7; DB 2; Length 1026;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 QTIGKRP 134
Db 861 QTIGKRP 867
```

```
RESULT 6
US-08-322-760A-6
; Sequence 6, Application US/08322760A
; Patent No. 5877279
; GENERAL INFORMATION:
; APPLICANT: Goldberg, Edward B.
; TITLE OF INVENTION: MATERIALS FOR THE PRODUCTION OF
; TITLE OF INVENTION: NANOMETER STRUCTURES AND USE THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie and Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/322,760A
; FILING DATE: 13-OCT-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 8471-0003-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: 212-869-8864
; TELEX: 66441 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1026 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Bacteriophage T4
; IMMEDIATE SOURCE:
; CLONE: p37 amino acid
; US-08-322-760A-6

Query Match          1.6%; Score 7; DB 2; Length 1026;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 QTIGKRP 134
Db 861 QTIGKRP 867

RESULT 7
US-08-974-775-30
; Sequence 30, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775
FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-30

Query Match 1.4%; Score 6; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
Db 5 NHEATS 10

RESULT 8
US-08-974-775-29
Sequence 29, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
APPLICANT: Toback, F. Gary
APPLICANT: Walsh-Reitz, Margaret
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
STREET: Plaza Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775
FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid

STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-29

Query Match 1.4%; Score 6; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
Db 5 NHEATS 10

RESULT 9
US-08-974-775-28
Sequence 28, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
APPLICANT: Toback, F. Gary
APPLICANT: Walsh-Reitz, Margaret
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
STREET: Plaza Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775
FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-28

Query Match 1.4%; Score 6; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
Db 5 NHEATS 10

RESULT 10
US-08-974-775-8
Sequence 8, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:

; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-775-8

Query Match 1.4%; Score 6; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
Db 8 NHEATS 13

RESULT 11
US-08-974-775-27
; Sequence 27, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775

; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-775-27

Query Match 1.4%; Score 6; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
Db 5 NHEATS 10

RESULT 12
US-08-974-775-9
; Sequence 9, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-775-9

Query Match 1.4%; Score 6; DB 3; Length 14;

Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
Db 8 NHEATS 13

RESULT 13

US-08-974-775-10
; Sequence 10, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-974-775-10

Query Match 1.4%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
Db 8 NHEATS 13

RESULT 14

US-08-974-775-32
; Sequence 32, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE

; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 14
; OTHER INFORMATION: /product= "Ser or Tyr"
US-08-974-775-32

Query Match 1.4%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
Db 8 NHEATS 13

RESULT 15

US-08-974-775-5
; Sequence 5, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:

```

; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-974-775-5

```

```

Query Match      1.4%; Score 6; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
    |||||
Db 8 NHEATS 13

```

Search completed: May 9, 2002, 12:55:42
Job time: 552 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:14:18 ; Search time 44.45 Seconds
(without alignments)
757.461 Million cell updates/sec

Title: US-09-787-083-2

Perfect score: 2360

Sequence: 1 MKVSLSTLTSLSCFAILA.....YNHEATSFVGLMLNDWMGL 442

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	814	34.5	382	2 E81195	phospholipase A1,
2	814	34.5	409	2 H81831	probable phospholip
3	373	15.8	286	2 B36971	outer membrane pho
4	367	15.6	329	2 D81279	phospholipase A1 (
5	358	15.2	289	2 A36971	outer membrane pho
6	357	15.1	289	1 PSECA1	phospholipase A1 (
7	357	15.1	289	2 E86069	outer membrane pho
8	321.5	13.6	289	2 C36971	outer membrane pho
9	246.5	10.4	355	2 C64582	phospholipase A1 p
10	240.5	10.2	355	2 H71930	probable phospholi
11	112.5	4.8	602	1 TVRTRR	protein kinase (EC
12	110.5	4.7	800	2 A29003	cellulase (EC 3.2.
13	110.5	4.7	822	2 JT0611	cellulase (EC 3.2.
14	109	4.6	824	3 JC7532	cellulase (EC 3.2.
15	107.5	4.6	901	2 T20122	hypothetical prote
16	104	4.4	783	2 JC5467	cellulase (EC 3.2.
17	101.5	4.3	660	2 E83656	methionyl-tRNA syn
18	101.5	4.3	838	2 A96557	probable receptor
19	100.5	4.3	1658	2 T42642	phosphoinositide 3
20	99	4.2	5005	2 F82884	hypothetical prote
21	98.5	4.2	797	2 T46737	X-Pro dipeptidyl-p
22	98.5	4.2	810	2 S49744	AMP deaminase (EC
23	97	4.1	4307	2 T20721	hypothetical prote
24	96	4.1	719	2 A42893	penicillin-binding
25	96	4.1	888	2 S50801	AMP deaminase homo
26	95	4.0	791	2 H96839	hypothetical prote
27	94.5	4.0	397	2 A35136	cellulase (EC 3.2.
28	94.5	4.0	564	2 T40777	ferric reductase t
29	94.5	4.0	601	2 T26062	hypothetical prote

ALIGNMENTS

RESULT 1

E81195

phospholipase A1, probable NMB0464 [imported] - Neisseria meningitidis (strain MC58 s

C;Species: Neisseria meningitidis

C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C;Accession: E81195

R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.

ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.

Science 287, 1809-1815, 2000

A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;

A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A;Reference number: A81000; MUID:20175755

A;Accession: E81195

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-382 <TET>

A;Cross-references: GB:AE002403; GB:AE002098; NID:g7225688; PIDN:AAF40901.1; PID:g722

A;Experimental source: serogroup B, strain MC58

C;Genetics:

C;Gene: NMB0464

Query Match 34.5% Score 814; DB 2; Length 382;
Best Local Similarity 44.4%; Pred. No. 3.6e-58;
Matches 162; Conservative 68; Mismatches 111; Indels 24; Gaps 8;

Qy 87 INCNALNODIMRLACYDTLVHGETPAVI-----RTKRSIRLDETWTI-RGKPVVIYQE 140

Db 33 LQCAALTDNVTRLACYDRIFAAQLPSSAGQESQESKAVLNLTEIVRSSLDKGEAVIVVEK 92

Qy 141 TTDPIFLMGNEKGMMLTKDAKQLEYAAKQFTPLSLFSDLDNN-TPLWSSRPHNPMYVLP 199

Db 93 GGDAL-----PADSAGETADITYPLSLMYDLKNDLRGLLVGVRHNPMLMP 139

Qy 200 IFMHGKNRSPNTPSH-EAKQFTNEFRAPELKFQVYKVKAAEDLWCTDSDLNFGYTTQ 258

Db 140 LWNNSNPAPGSPTRGTTVOEKFGQOKRAETKLVQVSKIAEDLFTRADLFGYITQR 199

Qy 259 SHWQIFN-GKNSRPFRRVHDYQPEIFLTQPVYSDLPWDCVKRMIGMGAHVHSHNGESAKLSR 317

Db 200 SDWQIYNQGRKSAPRNTDYKPEIFLTQPVKADLPFGGLRMLGAFVHQSGQSRPSR 259

Qy 318 SWNRAYLMAGMEWKNLTVMPIRWIRIFKSGSQPDNDPDLIDYVYGVDVRFPLENKS 377

Db 260 SWNRAYLMAGMEWKNLTVMPIRWVRAFTQ-SGDK-NDNPDIAVMGYGDVRLQYRLNDRQ 317

Qy 378 NISGTVRYNPRSGKALQLDVYVPLGKIGISGYFQIFQYQSGSLIDYNHEATSFVGLMLN 437

Db 318 NVYSVLRYNPKTGTGATEAAAYTFPIKGLKGVGRFGHGBESLIDYNHNKQNGIGIGLMFN 377

Qy 438 DWMGL 442

hypothetical prote
probable involveme
penicillin-binding
penicillin-binding
penicillin-binding
chitinase - Strept
protein p120 - Myc
protein kinase BCK
penicillin-binding
penicillin-binding
S-receptor kinase
hypothetical prote
cog-2 protein - Ca
hypothetical prote
involucrin - mouse
gene 14 protein -
hypothetical prote

30 94.5 4.0 655 2 T26061
31 94.5 4.0 1310 2 T40135
32 94 4.0 719 2 S28031
33 94 4.0 719 2 S28033
34 94 4.0 765 2 T35719
35 94 4.0 1078 2 T18352
36 93.5 4.0 1478 2 S20117
37 93 3.9 719 2 S28034
38 93 3.9 719 2 S28032
39 92.5 3.9 857 1 A41369
40 92 3.9 324 2 B69321
41 92 3.9 470 2 T43675
42 92 3.9 1788 2 T29043
43 91.5 3.9 467 1 A49377
44 91.5 3.9 486 2 S30959
45 91.5 3.9 711 2 E71673

Qy 422 DYNHEATSFVGLMLN 437
 Db 313 DYNKHLQLSTGFLIS 328

RESULT 5
 A36971
 outer membrane phospholipase A (EC 3.1.1.-) precursor - Salmonella typhimurium
 C:Species: Salmonella typhimurium
 C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 18-Jun-1999
 C:Accession: A36971; S40131
 R:Brok, R.G.P.M.; Brinkman, E.; van Bortel, R.; Bekkers, A.C.A.P.A.; Verheij, H.M.; Tom
 J. Bacteriol. 176, 861-870, 1994
 A:Title: Molecular characterization of enterobacterial pida genes encoding outer membra
 A:Reference number: A36971; MUID:94131966
 A:Accession: A36971
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-289 <BRO>
 A:Cross-references: GB:X76900; NID:g437024; PIDN:CAA54222.1; PID:g437025
 C:Genetics:
 A:Gene: pida
 A:Superfamily: bacterial phospholipase A1
 C:Keywords: carboxylic ester hydrolase

Query Match 15.2%; Score 358; DB 2; Length 289;
 Best Local Similarity 39.3%; Pred. No. 2e-21;
 Matches 84; Conservative 30; Mismatches 90; Indels 10; Gaps 4;

Qy 226 RAPELKFQSVKVAEDLWGTDSDLMFGYTOQSHWIFNGKNSRPRVHDYQPEIFLTQ 285
 Db 83 RKDEVKFLQSLAPFLWRGILGNPSVLGASYTKQSWQLSNSKSSPPRETNYEPQLFLGF 142

Qy 286 PV-YSDLPWDGKVRMIGMGAHVHNSGESAKLSRSNRAYLMAGMEWKNLTVMPRIWGRIF 344
 Db 143 ATDYRFAGW--TLRHVEMGYNHDNSGRSDPTSRTNRLYTLMAENGWNLVEVKPWVI- 199

Qy 345 KEGSGSQPDNDPDLTYGYGDRVFLYQLENKSNISGTVNPRSGKGLQLDYVYPLGK 404
 Db 200 -----GSTDNPDITKYGYYQLKIGYHL-GEAVLSAKGYNNWNTGYGGAELGSLYPVTK 253

Qy 405 GISGYFOIFGYGQSLIDYNHEATSFVGLMLND 438
 Db 254 HVRLYTQVYSGYGESLIDYNFNQTRGVGVMLND 287

RESULT 6
 PSECA1
 phospholipase A1 (EC 3.1.1.32) precursor - Escherichia coli
 N:Alternate names: outer-membrane phospholipase A; phosphatidylcholine 1-acylhydrolase;
 C:Species: Escherichia coli
 C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 11-Jun-1999
 C:Accession: A22133; S30711; A00771; S66447; F65186
 R:Homma, H.; Kobayashi, T.; Chiba, N.; Karasawa, K.; Mizushima, H.; Kudo, I.; Inoue, K.;
 J. Biochem. 96, 1655-1664, 1994
 A:Title: The DNA sequence encoding pida gene, the structural gene for detergent-resistan
 A:Reference number: A22133; MUID:85157492
 A:Accession: A22133
 A:Molecule type: DNA
 A:Residues: 1-289 <HOM>
 A:Cross-references: GB:X02143; GB:X00780; NID:g42423; PIDN:CAA26081.1; PID:g757840
 A:Note: this enzyme is tightly bound to the outer membrane of the cell
 R:Daniels, D.L.; Plunkett III, G.; Burland, V.; Blattner, F.R.
 Science 257, 771-778, 1992
 A:Title: Analysis of the Escherichia coli genome: DNA sequence of the region from 84.5
 A:Reference number: S30660; MUID:92358234
 A:Accession: S30711
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-13, 'FA', 16-289 <DAN>
 A:Cross-references: EMBL:M87049; NID:g836656; PIDN:AAA67617.1; PID:g148220

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1992
 R:de Geus, P.; Verheij, H.M.; Riegman, N.H.; Hoekstra, W.P.M.; de Haas, G.H.
 EMBO J. 3, 1799-1802, 1984
 A:Title: The pro- and mature forms of the E. coli K-12 outer membrane phospholipase A
 A:Reference number: A00771; MUID:85003590
 A:Accession: A00771
 A:Molecule type: DNA
 A:Residues: 'MTRQ', 34-289 <DEG>
 R:Dekker, N.; Merck, K.; Tommassen, J.; Verheij, H.M.
 Eur. J. Biochem. 232, 214-219, 1995
 A:Title: In vitro folding of Escherichia coli outer-membrane phospholipase A.
 A:Reference number: S66447; MUID:96048049
 A:Accession: S66447
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 21-24 <DEK>
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617
 A:Accession: F65186
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-289 <BLAT>
 A:Cross-references: GB:AE000458; GB:U00096; NID:g2367299; PIDN:AAC76824.1; PID:g23673
 A:Experimental source: strain K-12, substrain MG1655
 C:Comment: This enzyme is tightly bound to the outer membrane of the cell.
 C:Genetics:
 A:Gene: pida
 A:Map position: 85 min
 C:Superfamily: bacterial phospholipase A1
 C:Keywords: carboxylic ester hydrolase; membrane bound
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-289/Product: phospholipase A1 #status predicted <MPT>

Query Match 15.1%; Score 357; DB 1; Length 289;
 Best Local Similarity 39.3%; Pred. No. 2.4e-21;
 Matches 84; Conservative 29; Mismatches 91; Indels 10; Gaps 4;

Qy 226 RAPELKFQSVKVAEDLWGTDSDLMFGYTOQSHWIFNGKNSRPRVHDYQPEIFLTQ 285
 Db 83 RKDEVKFLQSLAPFLWRGILGNPSVLGASYTKQSWQLSNSKSSPPRETNYEPQLFLGF 142

Qy 286 PV-YSDLPWDGKVRMIGMGAHVHNSGESAKLSRSNRAYLMAGMEWKNLTVMPRIWGRIF 344
 Db 143 ATDYRFAGW--TLRDVEMGYNHDNSGRSDPTSRSNRLYTLMAENGWNLVEVKPWVYV- 199

Qy 345 KEGSGSQPDNDPDLTYGYGDRVFLYQLENKSNISGTVNPRSGKGLQLDYVYPLGK 404
 Db 200 -----GWTDNDPDTKYGYYQLKIGYHL-GDAVLSAKGYNNWNTGYGGAELGSLYPITK 253

Qy 405 GISGYFOIFGYGQSLIDYNHEATSFVGLMLND 438
 Db 254 HVRLYTQVYSGYGESLIDYNFNQTRGVGVMLND 287

RESULT 7
 E86069
 outer membrane phospholipase A [imported] - Escherichia coli (strain O157:H7)
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001
 C:Accession: E86069
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
 iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: E86069
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-289 <STO>

Qy 418 QSLIDYNHEATSFVGVLMLND 438
:||||: | :|:||||
Db 267 ESLIDYDFNQTRVGGMVMLND 287


```

Query Match          10.2%; Score 240.5; DB 2; Length 355;
Best Local Similarity 25.1%; Pred. No. 8.6e-12;
Matches 83; Conservative 44; Mismatches 109; Indels 95; Gaps 13;

QY   157 KKDAKOLEYAAKQTPLSLSFDLDRNNTPLWSSRPHNPMVYLPIFMHGKNRSPTPSHE 216
      || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB    69 KKYLNMDYGLTYFLPFYHSF-----TFIQWHYNINP----- 102

QY   217 AKQFTPNREFAPELKFQVSVKYKAAEDLWGTDSDLWFEGYTQQSHWOIFNGKNRPFRVHD 276
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB    103 ---YORNEF-----KFQISFRVPVFERHLMTKTGLYLAYTQTNNFQIYNDPQSAPMRMIN 154

QY   277 YQPEILFTPVSVDLPWDGKV---RMTCMGAVHHNSG-ESAKLSRSNR----- 321
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB    155 FMPLELIYYIPT-NFKPRGGKIGNSEIWIGHQHLSNGVGGAQCYPFNKEGNPENQPFGQ 213

QY   322 -----AYLMAG-----MEWK--NLVTMPRIWGRIFKEGSGSQDD- 354
      || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB    214 PVIIVKDNGQKDVWRGCRSVSAGNALCFVLVBEGGLKIMVAYPYV-----PYDQ 265

QY   355 -NPDIIDYGYGDVRFLY-----QLENKSNIISCTVRYNPRSGKGALQLDYVYPLGKGI 406
      || :||| ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB    266 SNPQLIDMGMGVNAKIDYRCGRHHFELQLYDIFYQYWRYD--RWHGAFRLCYTRYINPFV 323

QY   407 SGYQFIQFGYGQSLIDYNHEATSGVGGLMLN 437
      || :||| ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB    324 GIYAQWENGYGGLYEYDFVFSNRIGVGIRLN 354


RESULT 11
TVTRRR
protein kinase (EC 2.7.1.37) raf - rat
N:Alternate names: kinase-related transforming protein raf; raf proto-oncogene protein-S-
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 11-Jun-1999
C:Accession: B26126
R:Ishikawa, F.; Takaku, F.; Nagao, M.; Sugimura, T.
Mol. Cell. Biol. 7, 1226-1232, 1987
A>Title: Rat c-raf oncogene activation by a rearrangement that produces a fused protein.
A:Reference number: A26126; MUID:87172791
A:Accession: B26126
A:Molecule type: mRNA
A:Residues: 1-602 <ISH>
A:Cross-references: GB:M15428; NID:g206346; PIDN:AAA42002.1; PID:g206347
C:Genetics:
A:Gene: raf
C:Superfamily: rat protein kinase raf; protein kinase homology
C:Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; serine/threonin
F:301-567/Domain: protein kinase homology <XIN>
F:309-317/Region: protein kinase ATP-binding motif
F:329/Active site: Lys #status predicted
F:453/Binding site: phosphate (Ser) (covalent) (by autophosphorylation) #status predicted

Query Match          4.8%; Score 112.5; DB 1; Length 602;
Best Local Similarity 19.3%; Pred. No. 0.43;
Matches 93; Conservative 71; Mismatches 174; Indels 145; Gaps 23;

QY   23 QAQAVPPVPAVFDEVRENDLGQNDELPIDVQSATQASDTDTANPLDEHPELYTTALEN 82
      :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB    81 ELEKLNNOVKVLTTEKKNELETAQDRNLGI-----OSQFTRAKEELEAEKRDLIRT---N 131

QY   83 KTMLINCSALMODIMKLACVDYTLVGHGTTPAVIKTRILRDETIIQ--TTKGRQPVIYQE 140
      :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB    132 ERLSQVEYLTEDVKRL--NEKLKESNT---TKGELOLKLDLEOASDVTVKYREKRLQE 186

QY   141 -----TTDPIFLMGNEG-----MLTKDKADKOLEYAAKQFTPLS 174
      || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB    187 KELLHNQNSWLNETLTKTKTDALLGREKNGEILLELACTLENKKEADAIRSHSESASPASA 246

QY   175 LSFDLRDNNTPLWSSRPHNPMVYLPIFMHGKNRSPTPSHEAKQFTPNFEAPELKFQV 234
      || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB    247 LS-SSPNLSPGWSQPKTP-----VPAQREAPGSQTQEKNKIRPGORDSSSYWEI 298

```


R:McMurray, A.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z19225
A:Accession: T20122
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-901 <WIL>
A:Cross-references: EMBL:Z70750; PIDN:CAA94750.1; GSPDB:GN00023; CESP:F25B3.1
A:Experimental source: clone C50F4
R:Gardner, A.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z19406
A:Accession: T21324
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-901 <W12>
A:Cross-references: EMBL:Z70752; PIDN:CAA94758.1; GSPDB:GN00023; CESP:F25B3.1
A:Experimental source: clone F25B3
C:Genetics:
A:Gene: CESP:F25B3.1
A:Map position: 5
A:Introns: 25/3; 47/2; 199/1; 242/1; 315/2; 544/3; 600/1; 626/1; 646/1; 666/3; 706/3; 75

Query Match 4.6%; Score 107.5; DB 2; Length 901;
Best Local Similarity 19.8%; Pred. No. 2;
Matches 75; Conservative 60; Mismatches 171; Indels 73; Gaps 16;
QY 31 VAFVDEVSRSENDLGO-----DNELPIDVQSATQASDTTANPLDEHEPELYTTALENK 83
DB 468 VAMITEIRNOKLEEAVYRHIPDEVEVTPQLASRNPALNQPTDADDDDAETSNMRFERS 527
QY 84 TMLI-----NCSALNQDIMRLACYDTLVHGETPAVITKRSIRLDETIWOT 129
DB 528 NVSITHVTPCVGAIRASNRASPSKRDLELRQARDLIEKSTTPAATPSNRKASDEERRRE 587
QY 130 I-----KGKQVIYQETTDPIFLMGNEKGMTKKDAKOLEYAAKQF-----TPLSLSF 178
DB 588 VRRMLNEKHQHTAIPSTSSSPYPTERRIDGSNT--DLRRIELDVHFKKRDPSPTLVKQ 645
QY 179 LDRNNTPLSSRPHNPYVLPFPMHGKPNRSPN--TPS-----HEAKQTPNEFRAPEL 230
DB 646 YDPNDT-----PH-----VPAIGRTNGNRNDSSTPSSASTFDRVKRY--GSMRS 692
QY 231 KFOVSVKVK-----AEDLMGTDSDLWFGYTQ--SHWQ-----IFNGKNSRPFVRVHDYQ 280
DB 693 KESLQLMKQYGYMGNDFFESSQDALATPKFSSQWEKDVDDVEGTANELYRIDERISD 752
QY 281 IFLTQPVYSDLPWDGKVRMIGMGAVHHSNGESAKLSRSWNRAYLMAGMEWKN-LTYMPRI 339
DB 753 ITAQADVQD-----KIRETEVGSSEEMLTASYLELTNERNTLVHRQEVYNIETIROV 807
QY 340 WGRIFKEGS--GSQPDNDP 356
DB 808 TSEIDQLGKQINEVPDDFP 826

Search completed: November 30, 2001, 14:18:32
Job time: 254 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:17:53 ; Search time 28.02 Seconds
(without alignments)
578.367 Million cell updates/sec

Title: US-09-787-083-2
Perfect score: 2360
Sequence: 1 MKVSLSTLTLSTLSFCAYLA.....YNHEATSFVGLMLNDWMGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 segs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	373	15.8	286	1	PAL_KLEPN
2	358	15.2	289	1	PAL_SALTY
3	357	15.1	289	1	PAL_ECOLI
4	321.5	13.6	289	1	PAL_PROVU
5	110.5	4.7	800	1	GUN_BACS1
6	98.5	4.2	810	1	ADMD_YEAST
7	96	4.1	888	1	XJH0_YEAST
8	94.5	4.0	397	1	GUN_PAEPO
9	93.5	4.0	1478	1	BCK1_YEAST
10	93	3.9	716	1	BAC2_MOUSE
11	92.5	3.9	849	1	SRK6_BRAOL
12	92	3.9	324	1	YL70_ARCFU
13	92	3.9	1158	1	RI14_HUMAN
14	92	3.9	1788	1	YP72_CAEEL
15	91.5	3.9	467	1	INVO_MOUSE
16	91.5	3.9	486	1	VG14_BPML5
17	91.5	3.9	1024	1	Y075_MVCGE
18	90.5	3.8	379	1	XJHT_HAEIN
19	90	3.8	496	1	CATA_DICDI
20	90	3.8	719	1	BPBA_STRPN
21	89.5	3.8	503	1	C44_RAT
22	89.5	3.8	825	1	GUN3_BACS4
23	89	3.8	521	1	NPRE_BACAM
24	88.5	3.8	1379	1	MET_MOUSE
25	88	3.7	1131	1	PMAL_DUNBI
26	87.5	3.7	353	1	DCUP_BACSU
27	87.5	3.7	478	1	DHGB_ACTICA
28	87.5	3.7	669	1	AMY_ALTHA
29	87.5	3.7	794	1	SEIL_HUMAN
30	87.5	3.7	969	1	SACE_STRSL
31	87	3.7	483	1	MURE_CHLTR
32	87	3.7	537	1	P4H2_MOUSE
33	87	3.7	560	1	DTXH_CORBE

34	87	3.7	842	1	LPFC_SALTY
35	87	3.7	1115	1	DP3A_BACSU
36	87	3.7	1474	1	A2MG_HUMAN
37	86.5	3.7	1374	1	YC9A_SCHPO
38	86	3.6	491	1	TY3H_PHASP
39	86	3.6	666	1	PD14_RAT
40	86	3.6	828	1	BGAL_BRAOL
41	86	3.6	1627	1	ADPI_MYCPN
42	85.5	3.6	742	1	CD44_HUMAN
43	85.5	3.6	790	1	SEIL_MOUSE
44	85	3.6	417	1	DHMH_PARDE
45	85	3.6	517	1	EAS_DROME

ALIGNMENTS

RESULT 1	
PAL_KLEPN	
ID	PAL_KLEPN
AC	P37446; STANDARD; PRT; 286 AA.
DT	01-OCT-1994 (Rel. 30, Created)
DT	01-OCT-1994 (Rel. 30, Last sequence update)
DE	PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT
DE	PHOSPHOLIPASE A) (DR-PROPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-
DE	ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).
GN	PIDA.
OS	Klebsiella pneumoniae.
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC	Klebsiella.
OX	NCBI_TaxID=573;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=94131966; Pubmed=8300539;
RA	Brok R.G.P.M.; Brinkman E.; van Bostel R.; Bekkers A.C.A.P.,
RA	Verheij H.M.; Tommassen J.;
RT	"Molecular characterization of enterobacterial pida genes encoding
RT	outer membrane phospholipase A.,"
RL	J. Bacteriol. 176:861-870(1994).
CC	-1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE
CC	A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
CC	-1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O -> 1-ACYLGLYCERO-
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.
CC	-1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O -> 2-ACYLGLYCERO-
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.
CC	-1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
CC	-1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
CC	LOCATED THERE.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X76901; CAA54223.1; -
DR PIR; B36971; B36971.
DR PIR; S40129; S40129.
DR InterPro; IPR003187; PLAL.
DR Pfam; PF02253; PLAL; 1.
DR Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium.
FT SIGNAL 1 20 BY SIMILARITY.
FT CHAIN 21 286 PHOSPHOLIPASE A1.
FT ACT_SITE 161 161 BY SIMILARITY.
SQ SEQUENCE 286 AA; 32544 MW; 3E39F863085108A3 CRC64;

Query Match 15.8%; Score 373; DB 1; Length 286;
Best Local Similarity 42.1%; Pred. No. 8.1e-23;
Matches 90; Conservative 27; Mismatches 87; Indels 10; Gaps 5;

RL Nature 409:529-533(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic *Escherichia coli*
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [6]
RP SEQUENCE OF 30-289 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=85003590; PubMed=6383820;
RA de Geus P., Verheij H.M., Riegman N.H., Hoekstra W.P.M., de Haas G.H.;
RA "The pro- and mature forms of the *E. coli* K-12 outer membrane
RT phospholipase A are identical.";
RL EMBO J. 3:1799-1802(1984).
RN [7]
RP SEQUENCE OF 174-289 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=87115164; PubMed=3027506;
RA Irino N., Nakayama K., Nakayama H.;
RA "The recQ gene of *Escherichia coli* K12: primary structure and
RT evidence for SOS regulation.";
RL Mol. Gen. Genet. 205:298-304(1986).
RN [8]
RP MUTAGENESIS OF SER-172.
RX MEDLINE=94131966; PubMed=8300539;
RA Brok R.G.P.M., Brinkman E., van Bostel R., Bekkers A.C.A.P.,
RA Verheij H.M., Tommassen J.;
RA "Molecular characterization of enterobacterial plid genes encoding
RT outer membrane phospholipase A.";
RL J. Bacteriol. 176:861-870(1994).
RN [9]
RP ACTIVE SITE SER-164, AND PARTIAL SEQUENCE.
RX MEDLINE=91249806; PubMed=2040286;
RA Horrevorts A.J.G., Verheij H.M., de Haas G.H.;
RA "Inactivation of *Escherichia coli* outer-membrane phospholipase A by
RT the affinity label hexadecanesulfonyl fluoride. Evidence for an
RT active-site serine.";
RL Eur. J. Biochem. 198:247-253(1991).
CC A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
CC REQUIRED FOR EFFICIENT SECRETION OF BACTERIOICINS, SEEMS TO BE
CC DORMANT IN NORMAL GROWING CELLS.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- COPACITOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
CC LOCATED THERE.
CC -1- INDUCTION: BY MEMBRANE DAMAGING, FOR EXAMPLE, BY PHAGE-INDUCED
CC LYSIS OR TEMPERATURE SHOCK.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02143; CAA26081.1; -;
DR EMBL; M87049; AAA67617.1; -;
DR EMBL; AE000458; AAC76824.1; -;
DR EMBL; AE005613; AAC59017.1; -;
DR EMBL; AP002567; BAB38174.1; -;
DR EMBL; M30198; AAA24516.1; -;
PIR; A00771; PSECA.

DR PIR; A22133; PSECA1.
DR PIR; S30711; S30711.
DR Ecogene; EG10738; pldA.
DR InterPro; IPR003187; PLAL.
DR Pfam; PF02253; PLAL; 1.
KW Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium;
KW Complete proteome.
FT SIGNAL 1 20
FT CHAIN 21 289
FT ACT_SITE 164 164
FT MUTAGEN 172 172
FT CONFLICT 14 15
FT CONFLICT 30 33
SQ SEQUENCE 289 AA; 33163 MW; A688AD32AA60F218 CRC64;

Query Match 15.18; Score 357; DB 1; Length 289;
Best Local Similarity 39.38; Pred. No. 1.6e-21;
Matches 84; Conservative 29; Mismatches 91; Indels 10; Gaps 4;

QY 226 RAPELKFOVSVKVKAEDLWGTDSLWFGYTQOOSHWQIFNGKNSRPFVRVHDYQPEIFLTQ 285
DB 83 RKDEVKFSLSAFPLWRGILGPNLSVLGASYTKQSWQSLNSESSFFRETNYEPQLFLGF 142
QY 286 PV-YSDLPWDGKVRMIGMAVHHNGESAKLSRSNNRAYLMAGMEKMLTVMPIWGRIF 344
DB 143 ATDYRFAGW--TLRDVEMGYNHDSNGRSDPTSRSNRLYTRLMAENGWLVKVPWYV- 199
QY 345 KEGSGSQPDNDPDLIDYGYGDVRELYOLEKNSLNGISGVRYNPRSGKALQLDYVPLGK 404
DB 200 -----GNTDDNDPDTKMGYQLKTYGHL-GDAVLSAKGQVNMWNTGYGAELGLSPYTK 253
QY 405 GISGVFQIFOGYQSLIDYNEHTEATSGYGLMLND 438
DB 254 HVRLYQVYSGESLIDYNEHTEATSGYGLMLND 287

RESULT 4
PAL_PROVU
ID PAL_PROVU STANDARD; PRT; 289 AA.
AC P37447;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT
DE PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-
DE ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).
GN PLDA.
OS Proteus vulgaris.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Proteus.
OX NCBI_TaxID=585;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94131966; PubMed=8300539;
RA Brok R.G.P.M., Brinkman E., van Bostel R., Bekkers A.C.A.P.,
RA Verheij H.M., Tommassen J.;
RT "Molecular characterization of enterobacterial plid genes encoding
RT outer membrane phospholipase A.";
RL J. Bacteriol. 176:861-870(1994).
CC -1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE
CC A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- COPACITOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
CC LOCATED THERE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----


```

RESULT 8
GUN_PAERO
ID GUN_PAERO STANDARD; PRT: 397 AA.
AC P23348;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 15-JUL-1998 (Rel. 36, Last annotation update)
DE ENDOGLUCANASE (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE) (CELLULOSE).
OS Paenibacillus polymyxa (Bacillus polymyxa).
OC Bacteria; Firmicutes; Bacillus/clostridium group;
OC Bacillus/Staphylococcus group; Paenibacillus.
OX NCBI_TaxID=1406;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90170877; PubMed=2307659;
RA Baird S.D., Johnson D.A., Seligy V.L.;
RT "Molecular cloning, expression, and characterization of
RT endo-beta-1,4-glucanase genes from Bacillus polymyxa and Bacillus
RT circulans.";
RL J. Bacteriol. 172:1576-1586(1990).
CC -|- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
CC LINKAGES IN CELLULOSE.
CC -|- SIMILARITY: BELONGS TO CELLULOSE FAMILY A (FAMILY 5 OF GLYCOSYL
CC HYDROLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M33791; AAA22631.1; -
CC DR PIR: A35136; A35136.
CC DR HSSP: P54583; LECE.
CC DR InterPro: IPR001547; Glyco_hydro_F5.
CC DR Pfam: PF00150; cellulase_1.
CC DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; 1.
CC KW Cellulose degradation; Hydrolase; Glycosidase.
CC FT ACT_SITE 194 194 PROTON DONOR (BY SIMILARITY).
CC FT ACT_SITE 317 317 NUCLEOPHILE (BY SIMILARITY).
CC SQ SEQUENCE 397 AA; 44357 MW; B9C2E802C04F0A2A CRC64;
CC -----
Query Match 4.0%; Score 94.5; DB 1; Length 397;
Best Local Similarity 25.7%; Pred. No. 3.1;
Matches 46; Conservative 21; Mismatches 61; Indels 51; Gaps 10;
QY 301 GMGAVHNSGSAKLSRNNRAYLMAGMEKNTLVMPRIWGR-----IFKEG----- 347
Db 43 GNKIVDESKEAFAFNLN-----EGLETPNT-LHGLWSRMDMDLQVKKEGYNLIR 95
QY 348 -----SGSQPDD-----NPDIDYCYGVDVRELYQLENKSNISGTV-----RYNPRS 389
Db 96 LPYSNQLFDSSSRPDSIDYHKNPDLV---GLNPQIMDKLIEKAGQGQIILDRHRPS 152
QY 390 GKGAQLQDLYV--YPLGKIGSYGFIQFGY-----GSLIDYNHBEATSGVGLMLNDW 439
Db 153 G-QQSELWVTSQYPSRWISDKMLADRYKNNPTVIGADLHNEPHGQASWGVTGNASTDW 210
RESULT 9
BCK1_YEAST
ID BCK1_YEAST STANDARD; PRT: 1478 AA.
AC Q01389; P32894;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SERINE/THREONINE PROTEIN KINASE BCK1/SLK1/SSP31 (EC 2.7.-.-).
GN BCK1 OR SLK1 OR SSP31 OR LAS3 OR YJL095W OR J0906.

```

```

OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=921186847; PubMed=1545797;
RA Costigan C., Gehring S., Snyder M.;
RT "A synthetic lethal screen identifies SLK1, a novel protein kinase
RT homolog implicated in yeast cell morphogenesis and cell growth.";
RL Mol. Cell. Biol. 12:1162-1178(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92104496; PubMed=1840547;
RA Irie K., Araki H., Oshima Y.;
RT "A new protein kinase, SSP31, modulating the SMP3 gene-product
RT involved in plasmid maintenance in Saccharomyces cerevisiae.";
RL Gene 108:139-144(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=EG123;
RX MEDLINE=92107166; PubMed=1729597;
RA Lee K.S., Levin D.E.;
RT "Dominant mutations in a gene encoding a putative protein kinase
RT (BCK1) bypass the requirement for a Saccharomyces cerevisiae protein
RT kinase C homolog.";
RL Mol. Cell. Biol. 12:172-182(1992).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C;
RX MEDLINE=95176706; PubMed=7871887;
RA Miosga T., Boles E., Schaaff-Gerstenschlaeger I., Schmitt S.,
RA Zimmermann F.K.;
RT "Sequence and function analysis of a 9.74 kb fragment of
RT Saccharomyces cerevisiae chromosome X including the BCK1 gene.";
RL Yeast 10:1481-1488(1994).
RN [5]
RP SEQUENCE OF 602-1104 FROM N.A.
RA Cusick M.E.;
RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
CC -|- FUNCTION: SERINE/THREONINE PROTEIN KINASE INVOLVED IN A SIGNAL
CC TRANSDUCTION PATHWAY THAT PLAY A ROLE IN YEAST CELL MORPHOGENESIS
CC AND CELL GROWTH. THIS PATHWAY SEEMS TO STARTS BY SMP3; THEN
CC INVOLVE THE KINASE PRK1 THAT MAY ACT ON THIS KINASE. BCK1 PROBABLY
CC PHOSPHORYLATES MKK1 AND MKK2 WHICH THEMSELVES PHOSPHORYLATE THE
CC MPK1 KINASE.
CC -|- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
CC -|- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M84389; -; NOT_ANNOTATED_CDS.
CC DR EMBL: D10389; BAA01226.1; -
CC DR EMBL: X60227; CAA42788.1; -
CC DR EMBL: X77923; CAA54896.1; -
CC DR EMBL: Z49370; CAA89389.1; -
CC DR EMBL: Z49369; CAA89388.1; -
CC DR EMBL: M88604; AAA21179.1; -
CC DR PIR: S20117; S20117.
CC DR PIR: S22285; S22285.
CC DR PIR: JQ1118; JQ1118.
CC DR PIR: JQ1432; JQ1432.
CC DR HSSP: P24941; 1A01.
CC DR SGD; S0003631; BCK1.
CC InterPro: IPR000719; Euk_pkinase.
CC DR InterPro: IPR002290; Ser_thr_kin_actsite.

```

```
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
KW ATP-binding; Phosphorylation.
FT DOMAIN 1175 1440 PROTEIN KINASE.
FT NP_BIND 1181 1189 ATP (BY SIMILARITY).
FT BINDING 1204 1204 ATP (BY SIMILARITY).
FT ACT_SITE 1303 1303 BY SIMILARITY.
FT MOD_RES 1134 1134
FT VARIANT 1119 1119 T -> P (IN BCK1-19; ACTIVATION).
FT VARIANT 1120 1120 I -> K (IN BCK1-11; ACTIVATION).
FT VARIANT 1120 1120 I -> T (IN BCK1-16; ACTIVATION).
FT VARIANT 1146 1146 G -> V (IN BCK1-10; ACTIVATION).
FT VARIANT 1174 1174 A -> P (IN BCK1-20; ACTIVATION).
FT CONFLICT 59 59 F -> I (IN REF. 3).
FT CONFLICT 79 79 E -> V (IN REF. 2).
FT CONFLICT 264 264 A -> P (IN REF. 3).
FT CONFLICT 279 279 N -> I (IN REF. 3).
FT CONFLICT 703 714 RYPQTPSYVDR -> STPKPRVITME (IN REF. 3).
FT CONFLICT 795 795 S -> A (IN REF. 3).
FT CONFLICT 802 802 L -> V (IN REF. 3).
FT CONFLICT 808 808 A -> S (IN REF. 3).
FT CONFLICT 903 903 T -> N (IN REF. 3).
FT CONFLICT 919 919 T -> N (IN REF. 3).
FT CONFLICT 960 962 ADA -> RDR (IN REF. 5).
FT CONFLICT 1086 1104 RVPDPSSYEFIQDLNGK -> VPIAHTSYRMDLTVKIN
FT H (IN REF. 5).
SQ SEQUENCE 1478 AA; 164194 MW; D586C3A497A5BB33 CRC64;

Query Match 4.08; Score 93.5; DB 1; Length 1478;
Best Local Similarity 22.28; Pred. No. 23;
Matches 97; Conservative 52; Mismatches 156; Indels 131; Gaps 24;

QY 36 EVRSNDLQGNELPI---DVQSATQASATDTANPLDEHEPELYTTALENK--TWLINC 90
DB 950 DIFEENDITADAPMPDSDSDSDSDSDSDIWSKKKTAPE---TNNENKKDEKSDNS 1006
QY 91 ALNDQIMRLACYDLIVGETPAVITKTRIRLDETINQTKGKQVYQETDPIFLMGN 150
DB 1007 THSDEIF---YDS---QTQDKMERK-----MTRFSPVEVYQ-NLEKFFPRAN 1047
QY 151 EKGMLTKKDAKQLEVAKQETPLSLFDLDRNTPLWSSRPHNPWYVLPIMHGKPNRS- 209
DB 1048 LDKPIT---EGIASPTSPKSLDLSLLSPKNVASSRTEPSTPSRPVP-----POSSY 1094
QY 210 -----PNTPSHEAKQETPNFRA-----PELKFOYSVRKVAKE-DLWGTDSDLW 252
DB 1095 EFIDGLNGKNKPLNQAK--TPKRTKTIRTIAHEASLARKNSVKLRQNTKMWGT---RM 1149
QY 253 FGYTQSHWQIFNGKNRPRFVHDYQPEIFLTPQVYSDLPWDGKVRMIGMA-----V 305
DB 1150 VEVTEHNVISINKARNSK-----GEYKEFAW-MKGEIMKGSFGAVYICL 1193
QY 306 HHSNGESAKLSR-----SNRAYLMAGMEWKNLTVMPTWIRIFKEGSGSQDPDNDPI 358
DB 1194 NVTTGEMMAVQVEVPKYSSQNEAILS-----TV-----EALRSEVTLKOLDHLNI 1240
QY 359 LDYGYGVDVRLYOLGKNKSNISGTVRYNPRSGKALQIDYVPLGKISGVYFGQYGG 418
DB 1241 VQYLGK-----ENKNYI-----SLFLEYV--AGSGVGLIRMYGFEDE 1277
QY 419 SLIDYNHEATSGFVGL 434
DB 1278 PLI--KHLTTQVLKGL 1291

RESULT 10
BAC2_MOUSE STANDARD; PRT; 716 AA.
ID BAC2_MOUSE
```

```
AC PF07303;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE TRANSCRIPTION REGULATOR PROTEIN BACH2 (BTB AND CNC HOMOLOG 2).
GN BACH2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=BALB/C;
RX MEDLINE=97042438; PubMed=8887638;
RA Oyake T., Itoh K., Motohashi H., Hayashi N., Hoshino H., Nishizawa M.,
RA Yamamoto M., Igarashi K.;
RT "Bach proteins belong to a novel family of BTB-basic leucine zipper
RT transcription factors that interact with Mafk and regulate
RT transcription through the NF-E2 site.";
RL Mol. Cell. Biol. 16:6083-6095(1996).
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS AS REPRESSOR OR
CC ACTIVATOR, BINDS, IN-VITRO, TO NF-E2 BINDING SITES. PLAY IMPORTANT
CC ROLES IN COORDINATING TRANSCRIPTION ACTIVATION AND REPRESSION BY
CC MAFK.
CC -!- SUBUNIT: HETERODIMER OF BACH2 AND MAFK.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: EXPRESSION RESTRICTED TO MONOCYTES AND
CC NEURONAL CELLS.
CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY. CNC SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 1 BTB/POZ DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D85604; BAA13138.1; -
DR HSP; P05412; IFOS.
DR MGD; MGI:894679; Bach2.
DR InterPro; IPR000210; BTB_POZ.
DR InterPro; IPR001871; bZIP.
DR Pfam; PF00651; BTB; 1.
DR Pfam; PF00170; bZIP; 1.
DR SMART; SM00338; BRLZ; 1.
DR SMART; SM00225; BTB; 1.
DR PROSITE; PS00097; BTB; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
DR Transcription regulation; Activator; Repressor; DNA-binding;
DR Nuclear protein.
KW DOMAIN 37 103 BTB.
KW DOMAIN 162 168 POLY-GLU.
FT DNA_BIND 527 542 BASIC MOTIF.
FT DOMAIN 550 572 LEUCINE-ZIPPER.
SQ SEQUENCE 716 AA; 78935 MW; 9132B3731AE24333 CRC64;

Query Match 3.98; Score 93; DB 1; Length 716;
Best Local Similarity 22.88; Pred. No. 9.2;
Matches 100; Conservative 42; Mismatches 138; Indels 158; Gaps 24;

QY 14 SCFAILAQQAQVNP---VAFVDEV--RSEDLG-----ODNELPIDVQSATQAS 61
DB 124 SCFSFL---QTQLLNREDGLFVCKRDSACQRPQEDHGNASAGEEETMDSETARMCA 180
QY 62 TD-----TANPLDEH-----EPelyTTALENKTMLINCsalnQ-----DIMRLA 100
DB 181 TDQMLPDPISEFATAIPVAEKEALLPESEVPTDTKENSEK---GALTQIPRYKKYOLA 236
QY 101 C-----YDPLVHG-----ETPAVITKRSIRL-----DE 124
DB 101 C-----YDPLVHG-----ETPAVITKRSIRL-----DE 124
```



```

OC Archaeoglobus.
OX NCBI_TaxID=2234;
FN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klench H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kervag A.R., Graham D.E., Kyriades N.C.,
RA Fleischmann R.D., Quackenbush J., Lee H.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -1- SIMILARITY: CONTAINS 2 KELCH REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000955; AAB89093.1;
DR TIGR; AF2170;
DR InterPro; IPR001798; Kelch.
DR Pfam; PF01344; Kelch; 2.
KW Hypothetical protein; Repeat; Complete proteome.
FT REPEAT 229 276
FT REPEAT 277 323
FT REPEAT 324 360
FT SEQUENCE 324 AA; 36025 MW; C445388CFEB96B45 CRC64;
SQ
Query Match 3.9%; Score 92; DB 1; Length 324;
Best Local Similarity 20.1%; Pred. No. 3.7;
Matches 54; Conservative 44; Mismatches 72; Indels 98; Gaps 18;
QY 224 EFRAPELKFQVSVKAAED-----LWGTDSLDFGTYQSHWQIF-NGKNSRPRFV 275
DB 63 EFHPNP-RLEI---LLSSEDFSQREAADVGEILLIFGGT-----VFENGKYSPTDQIL 112
QY 276 DYQPEIFLTPVYSDLP-----W-DGKVRMI-----GMCAYVHNGESAKLSRS 318
DB 113 SFNPKLERLRVNASLPHPTSDVAAYVWGDSRVYIFLNNSECEVYAFYPSNESFAKLDSV 172
QY 319 -----W--NRAYL-----WAGWEKNLTVMPRIWGRIFKEGSGSQ 351
DB 173 CPHEPGCGVHVVWYGGKAYFFGEGVASFDPMGGFKW--IAFTDRVWVRAATVADG-- 228
QY 352 PDONPDILDYGYGDVRLYOLENKNISGT-----VYNPNSRGKALQLDVYVPLGKG-- 405
DB 229 -----YIFAIGSSGIAETKDEIIRFNPKTGE-LCEMTKLPVARGQA 270
QY 406 --ISG-YEQIF-----QGYQSLIDYNH 425
DB 271 VAVGGEYIYFGYTKDGYANEILLRYDY 298
RESULT 13
ID RI14_HUMAN STANDARD; PRT; 1158 AA.
AC P48552;
DT 01-FEB-1996 (Rel. 33, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE NUCLEAR FACTOR RI140 (NUCLEAR RECEPTOR INTERACTING PROTEIN 1).

```

```

GN NR1P1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast;
RX MEDLINE=95369246; PubMed=7641693;
RA Cavailles V., Dauvois S., L'Horset F., Lopez G., Hoare S.,
RA Kushner P.J., Parker M.G.;
RT "Nuclear factor RI140 modulates transcriptional activation by the
RT estrogen receptor."
RL EMBO J. 14:3741-3751(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21."
RL Nature 405:311-319(2000).
CC -1- FUNCTION: MODULATES TRANSCRIPTIONAL ACTIVATION BY THE ESTROGEN
CC RECEPTOR.
CC
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X84373; CAA59108.1;
DR EMBL; AL163207; CAB90396.1;
DR MIM; 602490;
KW Transcription regulation; Nuclear protein.
FT CONFLICT 124 124 P -> R (IN REF. 1).
FT CONFLICT 448 448 R -> G (IN REF. 1).
FT CONFLICT 721 726 NKGSE -> TGRVK (IN REF. 1).
FT CONFLICT 1080 1080 T -> A (IN REF. 1).
SQ SEQUENCE 1158 AA; 126941 MW; 81FC424968E9A5F6 CRC64;
Query Match 3.9%; Score 92; DB 1; Length 1158;
Best Local Similarity 19.3%; Pred. No. 21;
Matches 104; Conservative 70; Mismatches 150; Indels 214; Gaps 26;
QY 3 VSLSTLTLSCPAILAIQQAQVNPVAFVDEVRND-----LGDQNE 48
DB 458 VSLDNFTQSLNTW-----DPKVPDVKDKEDQDTSKNSKLNHSHQKVTLLQLLGHKN 510
QY 49 LPIDVQSATO-----SASTDTANPLDEHEPELYTTA-----L 80
DB 511 ENVEKNTSPQGVHNDVSKFNTQNTVARTSVIESPSTNRTTPVS--TPPLLTSSKAGSPINL 568
QY 81 ENKTMLIN-----CSALNQDINMLACYDTL-----VHCETPAVTK-----TKRSI 120
DB 569 SQHSLVIKWNPPVVCSTQSEKLTNTASNHSMDLTKSKDPPGKPAQNEGAQNSATFSAS 628
QY 121 RLDETITWTKGKPOVIYQETTDPIFLMGNEKMGKLTAKQKOLEYAAKQFTPLSLSDLD 180
DB 629 KLLQNLACQCGMQSSMSVEEQRPSSKOLLTGN-----TDK-----PIGM---ID 667

```


Qy 421 -----IDYNH 425
: : : :
Db 417 EKASREQQLDYSH 429

Search completed: November 30, 2001, 14:27:00
Job time: 547 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:17:03 ; Search time 77.9 seconds
(without alignments)
829.941 Million cell updates/sec

Title: US-09-787-083-2
Perfect score: 2360
Sequence: 1 MKVSLSTLTLSLSCFAILA.....YNHEATSGVGLMNDMMGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 14627329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- SPTREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	814	34.5	382	Q9K0U7	Q9K0U7 neisseria m
2	814	34.5	409	Q9J721	Q9J721 neisseria m
3	372	15.8	306	Q9CL22	Q9CL22 pasteurella
4	367	15.6	329	Q9PMU8	Q9PMU8 campylobact
5	360.5	15.3	292	Q9SID7	Q9SID7 yersinia ps
6	360	15.3	289	Q9L6N9	Q9L6N9 salmonella
7	356.5	15.1	292	Q9Z4N8	Q9Z4N8 enterobacte
8	342.5	14.5	297	Q32349	Q32349 campylobact
9	246.5	10.4	355	Q25241	Q25241 helicobacte
10	240.5	10.2	355	Q9ZLX5	Q9ZLX5 helicobacte
11	122	5.2	278	Q9XB53	Q9XB53 erwinia car
12	117	5.0	1686	Q00443	Q00443 homo sapien
13	112.5	4.8	602	Q63485	Q63485 rattus norv
14	110.5	4.7	821	Q59241	Q59241 bacillus sp
15	109	4.6	824	Q9F216	Q9F216 bacillus sp
16	107.5	4.6	901	Q18749	Q18749 caenorhabdi
17	105	4.4	435	Q9X9C0	Q9X9C0 streptococc
18	105	4.4	783	Q45554	Q45554 bacillus sp
19	103.5	4.4	1046	Q04941	Q04941 streptococc

20	103	4.4	798	4	Q9UN32	Q9UN32 homo sapien
21	103	4.4	798	4	Q9UBK2	Q9UBK2 homo sapien
22	102.5	4.3	3247	12	Q65553	Q65553 bovine herp
23	101.5	4.3	403	5	P91736	P91736 hydra magni
24	101.5	4.3	660	2	Q9KX8	Q9KX8 bacillus ha
25	101.5	4.3	739	5	Q9V9E6	Q9V9E6 drosophila
26	101.5	4.3	838	10	Q9C815	Q9C815 arabidopsis
27	100.5	4.3	1509	11	Q61194	Q61194 mus musculu
28	100.5	4.3	1658	11	Q61182	Q61182 mus musculu
29	99.5	4.2	788	10	Q9SCV5	Q9SCV5 arabidopsis
30	99	4.2	422	2	Q9RC26	Q9RC26 streptomyce
31	99	4.2	5005	2	Q9PPZ5	Q9PPZ5 ureaplasma
32	98.5	4.2	467	4	Q9UJM5	Q9UJM5 homo sapien
33	98.5	4.2	467	4	Q9BTY2	Q9BTY2 homo sapien
34	98.5	4.2	797	2	Q9RDW6	Q9RDW6 lactobacill
35	97.5	4.1	683	2	Q9A6U7	Q9A6U7 caulobacter
36	97	4.1	576	2	Q9L115	Q9L115 streptomyce
37	97	4.1	4307	5	Q19319	Q19319 caenorhabdi
38	96	4.1	871	4	Q9Y5C3	Q9Y5C3 homo sapien
39	96	4.1	938	4	Q9Y5F7	Q9Y5F7 homo sapien
40	95	4.0	690	5	Q61142	Q61142 plasmodium
41	95	4.0	719	2	Q57114	Q57114 streptococc
42	95	4.0	719	2	Q9RET8	Q9RET8 streptococc
43	95	4.0	791	10	Q49137	Q49137 arabidopsis
44	94.5	4.0	476	5	Q9BI60	Q9BI60 caenorhabdi
45	94.5	4.0	530	5	Q45879	Q45879 caenorhabdi

ALIGNMENTS

RESULT 1

ID	Q9K0U7	PRELIMINARY;	PRT;	382 AA.
AC	Q9K0U7;			
DT	01-OCT-2000 (TrEMBLrel. 15, Created)			
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)			
DE	PHOSPHOLIPASE A1, PUTATIVE.			
GN	NMB0464.			
OS	Neisseria meningitidis (serogroup B).			
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.			
OX	NCBI_TaxID=491;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MC58 / SEROGROUP B:			
RA	MEDLINE=20175755; PubMed=10710307;			
RA	Tetelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,			
RA	Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,			
RA	Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,			
RA	Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,			
RA	Mason T., Ciecko A., Parksey D.S., Blair E., Citterone H., Clark E.B.,			
RA	Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,			
RA	Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,			
RA	Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;			
RT	"Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.";			
RL	Science 287:1809-1815(2000).			
DR	EMBL; AE002403; AAF40901.1; -.			
DR	TIGR; NMB0464; -.			
DR	InterPro; IPR003187; PLAI.			
DR	Pfam; PF02253; PLAI; 1.			
KW	Complete proteome.			
SQ	SEQUENCE 382 AA; 42714 MW; B468A802F062E836 CRC64;			

Query Match 34.5%; Score 814; DB 2; Length 382;
Best Local Similarity 44.4%; Pred. No. 1e-62;
Matches 162; Conservative 68; Mismatches 111; Indels 24; Gaps 8;

Qy 87 INCSALNDIWRACLDYLVHGETPAVI-----KTKRSIRLDETWTQTI-KGKPOVIQVE 140

Db 33 LQCAALTDNVTRLACYDRIFAQLPSSAGQEQESKAVILNLTETVRSLLDKGEAVIVVEK 92

```
QY 141 TTDPIFLMGNEKGLTKKDAKOLEYAAKQFTPLSLFSDLDNRN--TPLWSSRPHNPMYVLP 199
Db 93 GGDAL-----PADSAGETADIVTPLSLMYDLKNDLRLGLLVREHNPYMLP 139
QY 200 IFMHGKPNRSPNTPSH-EAKQFTPNFRAPDELKQVSVKVAEDLWGTDSDLWFGYTQ 258
Db 140 LWTNNSPNYAPGSPTRGTTVQERFGQOKRAETKLOVSFKIAEDLFKTRADLWFGYTQ 199
QY 259 SHWOIEN-GKNSRPRVHDYQPEIFLTQPVYSDLPWDGKVRMIGMGAVHHSNGESAKLSR 317
Db 200 SDWOIYNOGRKSAFFRTWTDYKPEIFLTQPVKADLPFGGRMLRLMGAGFVHOSGQSRPESR 259
QY 318 SWNRAYLMAGMEWKNLTVMPRIMGRIFKEGSGQPDNDPILDYGYGQVRFYQLLENKS 377
Db 260 SWNRAYAMAGMEWKNLTVPVWVWVAFDQ--SGDK-NDNPDIAIDYMGYGVKQVLYRLNDRQ 317
QY 378 NISGTVRYNPRSGKALQOLDYVYPLGKIGISGYFOIFOGYGOSLIDYNHEATSFVGGLMLN 437
Db 318 NVYSVLRYNPKTYGAEAAATTPKIKGLKGVVGRFGHYGESLIDYNHKQNGIGIGLMFN 377
QY 438 DWMGL 442
Db 378 DLGDI 382
RESULT 2
Q9JUT21
ID Q9JUT21 PRELIMINARY; PRT; 409 AA.
AC Q9JUT21
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE PUTATIVE PHOPHOLIPASE.
GN NWA2021.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
Whithead S., Pratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis 22491."
RL Nature 404:502-506(2000).
DR EMBL: AL162757; CAB85240.1; -.
DR InterPro: IPR003187; PLAI.
DR Pfam: PF02253; PLAI; 1.
KW Complete proteome.
SQ SEQUENCE 409 AA; 45862 MW; CD65858064D01AA1 CRC64;
```

```
Query Match 34.5%; Score 814; DB 2; Length 409;
Best Local Similarity 44.4%; Pred. No. 1.2e-62;
Matches 162; Conservative 68; Mismatches 111; Indels 24; Gaps 8;
```

```
QY 87 INCSALNODIMRLACDYTLVHGTEPAVI-----KTKRSIRLDEFITQTI-KGKQVYIYQE 140
Db 60 LQCAALTDNVTRLACDYRIFAALPSAGQEGESKAVLNLTETVRSRLDKGEAVIVVEK 119
QY 141 TTDPIFLMGNEKGLTKKDAKOLEYAAKQFTPLSLFSDLDNRN--TPLWSSRPHNPMYVLP 199
Db 120 GGDAL-----PADSAGETADIVTPLSLMYDLKNDLRLGLLVREHNPYMLP 166
QY 200 IFMHGKPNRSPNTPSH-EAKQFTPNFRAPDELKQVSVKVAEDLWGTDSDLWFGYTQ 258
Db 167 LWTNNSPNYAPGSPTRGTTVQERFGQOKRAETKLOVSFKIAEDLFKTRADLWFGYTQ 226
```

```
QY 259 SHWOIEN-GKNSRPRVHDYQPEIFLTQPVYSDLPWDGKVRMIGMGAVHHSNGESAKLSR 317
Db 227 SDWOIYNOGRKSAFFRTWTDYKPEIFLTQPVKADLPFGGRMLRLMGAGFVHOSGQSRPESR 286
QY 318 SWNRAYLMAGMEWKNLTVMPRIMGRIFKEGSGQPDNDPILDYGYGQVRFYQLLENKS 377
Db 287 SWNRAYAMAGMEWKNLTVPVWVWVAFDQ--SGDK-NDNPDIAIDYMGYGVKQVLYRLNDRQ 344
QY 378 NISGTVRYNPRSGKALQOLDYVYPLGKIGISGYFOIFOGYGOSLIDYNHEATSFVGGLMLN 437
Db 345 NVYSVLRYNPKTYGAEAAATTPKIKGLKGVVGRFGHYGESLIDYNHKQNGIGIGLMFN 404
QY 438 DWMGL 442
Db 405 DLGDI 409
RESULT 3
Q9CL22
ID Q9CL22 PRELIMINARY; PRT; 306 AA.
AC Q9CL22;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE HYPOTHETICAL PROTEIN PM1426.
GN PM1426.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL: AE006179; AAK03510.1; -.
DR InterPro: IPR003187; PLAI.
DR Pfam: PF02253; PLAI; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 306 AA; 35580 MW; EAF3DE8C1C22B26E CRC64;
```

```
Query Match 15.8%; Score 372; DB 2; Length 306;
Best Local Similarity 39.0%; Pred. No. 2.5e-24;
Matches 83; Conservative 40; Mismatches 82; Indels 8; Gaps 4;
```

```
QY 229 ELKFOVSVKVKAAEDLWGTDSDLWFGYTQOSHWOIENKNSRPRVHDYQPEIFLTQPVY 288
Db 97 EIKFKISLALPLWRGILGNNSVLAASYTKSWFOLSNVDDSSPPRETNYEPQLFLANKTQ 156
QY 289 SDLPWDGKVRMIGMGAVHHSNG--ESAKLSRSMNRAYLMAGMEWKNLTVMPRIMGRIFKE 346
Db 157 YSLPFGWTLDQVETGINHQSGNRDADAEKLSNRNLVVRASAIKONNVTVEIKPMWRIPEK 216
QY 347 GSGSQPDNDPILDYGYGQVRF--LYOLENKSNTSGTVRYNPRSGKALQOLDYVYPLGKG 405
Db 217 ---AKNDNDPDITKYRGHFDVALGYHYHQQFKLSG--HYNPISNKGGLSEASYSPITKN 271
QY 406 ISGVFOIFOGYGOSLIDYNHEATSFVGGLMLND 438
Db 272 IRFTQYNYGYGESLIDYQRIORIGIGISLNN 304
```

```
RESULT 4
Q9PMU8
ID Q9PMU8 PRELIMINARY; PRT; 329 AA.
AC Q9PMU8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
```


RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Kaip P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen *Helicobacter*
RT *pylori*.";
RL Nature 388:539-547(1997).
DR EMBL; AE000564; AAD07564.1; -.
DR TIGR; HP0499; -.
DR InterPro; IPR003187; PLA1.
DR Pfam; PF02253; PLA1; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 355 AA; 42486 MW; 461960F15E65AB0A CRC64;

Query Match 10.4%; Score 246.5; DB 2; Length 355;
Best Local Similarity 25.1%; Pred. No. 2.6e-13;
Matches 82; Conservative 41; Mismatches 117; Indels 87; Gaps 11;
QY 157 KKDAKOLEYAAKQFTPLSLSFDRNTPLMSSRPHNPMYVLPFPMHGKPNRSPNTPSHE 216
DB 69 KKYLNMDYLGTYFLPFVHSF-----TFIQWYHPNINP----- 102
QY 217 AKQFTPNFRAPFELKFOYSVKVKAEDLWGTDSDLWFGYTOQSHWQIFNGKNSRPFVRVD 276
DB 103 ----YORNEF-----KQISFRVPVFRHILWTGTLYLAYTQTFNFOIYNDSAPMRMN 154
QY 277 YOPEIFLTQPVYSDLPWDGKV---RMIGMGAVHHSNG--ESAKLSRSNRLAYLMAGMEWKN 332
DB 155 FNPFLIYVYPI-NFKPFGKIGNFSEIWGWHSNGVGAOCYQPFNKEGPNQ 209
QY 333 LTVMPRI-----WGRIFKGGSGSQ-----DNDPD 357
DB 210 FPGQVIVKDYNGQKDVRRGGRSVSAGQRPVRLVWEKGGLKIMVAYWPVYDQSNPN 269
QY 358 ILDYGYGQDVRLY-----QLENKSNISGTVRNPBSGKALQLDVYPLGKISGYF 410
DB 270 LIDYMGYGNADYRGRHHELFQLYDFTQWRYD--RWHGAFRLGTYRINPFGIYA 327
QY 411 QIFQCGYQSLIDYNHEATSFVGVLMLN 437
DB 328 QWFNGYGDLYEYDFVFSNRIGVGIRLN 354

RESULT 10
Q9ZLX5 PRELIMINARY; PRT; 355 AA.
AC Q9ZLX5;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE PUTATIVE PHOSPHOLIPASE A1.
GN PLDA OR JHP0451.
OS *Helicobacter pylori* J99 (Campylobacter *pylori* J99).
OC Bacteria; Proteobacteria; epsilon subdivision; *Helicobacter* group;
OC *Helicobacter*.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923582;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Dolg P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen *Helicobacter pylori*.";
RL Nature 397:176-180(1999).
DR EMBL; AE001479; AAD06029.1; -.
DR InterPro; IPR003187; PLA1.
DR Pfam; PF02253; PLA1; 1.
KW Complete proteome.
SQ SEQUENCE 355 AA; 42329 MW; B3CE9810EBA2FDC9 CRC64;

Query Match 10.2%; Score 240.5; DB 2; Length 355;
Best Local Similarity 25.1%; Pred. No. 8.8e-13;
Matches 83; Conservative 44; Mismatches 109; Indels 95; Gaps 13;
QY 157 KKDAKOLEYAAKQFTPLSLSFDRNTPLMSSRPHNPMYVLPFPMHGKPNRSPNTPSHE 216
DB 69 KKYLNMDYLGTYFLPFVHSF-----TFIQWYHPNINP----- 102
QY 217 AKQFTPNFRAPFELKFOYSVKVKAEDLWGTDSDLWFGYTOQSHWQIFNGKNSRPFVRVD 276
DB 103 ----YORNEF-----KQISFRVPVFRHILWTGTLYLAYTQTFNFOIYNDSAPMRMN 154
QY 277 YOPEIFLTQPVYSDLPWDGKV---RMIGMGAVHHSNG--ESAKLSRSNRLAYLMAGMEWKN 321
DB 155 FNPFLIYVYPI-NFKPFGKIGNFSEIWGWHSNGVGAOCYQPFNKEGPNQ 213
QY 322 -----AYLMAG-----MEWK--NLTVMPRIWGRIFKGGSGSQPDD- 354
DB 214 PVIVKDYNGQKDVRRGGRSVSAGNALCFVLVWEKGGLKIMVAYWPV-----PYDQ 265
QY 355 -NPDILDYGYGQDVRLY-----QLENKSNISGTVRNPBSGKALQLDVYPLGKGI 406
DB 266 SNQIDYMGYGNADYRGRHHELFQLYDFTQWRYD--RWHGAFRLGTYRINPFGV 323
QY 407 SGYFQIFQCGYQSLIDYNHEATSFVGVLMLN 437
DB 324 GIYAQWFGYGDLYEYDFVFSNRIGVGIRLN 354
RESULT 11
Q9XB53 PRELIMINARY; PRT; 278 AA.
AC Q9XB53;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
DE KDU1.
GN KDU1.
OS *Erwinia carotovora*.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC *Pectobacterium*.
OX NCBI_TaxID=554;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 39048, GS101;
RX MEDLINE=98065591; PubMed=9402024;
RA McGowan S.J., Sebahia M., O'Leary S., Hardie K.R., Williams P.,
RA Stewart G.S., Bycroft B.W., Salmund G.P.;
RT "Analysis of the carbenem gene cluster of *Erwinia carotovora*:
RT definition of the antibiotic biosynthetic genes and evidence for a
RT novel beta-lactam resistance mechanism.";
RL Mol. Microbiol. 26:545-556(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 39048, GS101;
RX MEDLINE=98276484; PubMed=9614345;
RA McGowan S.J., Bycroft B.W., Salmund G.P.;
RT "Bacterial production of carbenems and clavams: evolution of beta-
RT lactam antibiotic pathways.";
RL Trends Microbiol. 6:203-208(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 39048, GS101;
RA McGowan S.J.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; U17224; AAD38237.1; -.
SQ SEQUENCE 278 AA; 31341 MW; 63769F4F3550E1B7 CRC64;

Query Match 5.2%; Score 122; DB 2; Length 278;
Best Local Similarity 21.3%; Pred. No. 0.013;
Matches 73; Conservative 53; Mismatches 128; Indels 88; Gaps 16;

```
Query Match 5.0%; Score 117; DB 4; Length 1686;
Best Local Similarity 19.2%; Pred. No. 0.48;
Matches 86; Conservative 67; Mismatches 154; Indels 140; Gaps 21;

QY 51 IDVQSATQASASTDTANPLDEHEPELYTTALENKTMLNCSALNODIMRLACYDTLVHGET 110
   :||: :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 1 MDVRSVHS-----EHAKTLDTELKFKELIQETPNQYNTWYSHIDRIVVG-- 48

QY 111 PAVIKTRSRRLDETIWQ-----TIKKGQVYIYQETDPIFLMGNEKGM 154
   :||: :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 49 -GIMPVGEITFDGIGKQGVNYFLERRELGLINIGGPAKIVIDGTS---YEVGNEAL 104

QY 155 LTKKDAKQLEYAAKQFTPLSLFDLDRNNTPLWSSRPHNPMYVLPFPMHCKPNRSPNTPS 214
   :||: :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 105 YVGGKAKALAFS-----SLDSAKPAKLYNSAPAHAFVPTRIITQDDAKAPLGDV 155

QY 215 HEAKQFTPNFRAPELKFOVSVKAAEDLWGTDSDLWFGYT---QOSHWOIENGKNSRP 271
   :||: :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 156 KTCNKRITICKLYPEV-----VETCQ-----LSMGLTRLAEGSNW-----NSMP 194

QY 272 FRVHYOPEIFLTPQVYSDLPWPKGVKRMIGAVHHS-----NGESAKLSRNSRAYLMA 326
   :||: :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 195 THTHERRMEVVF---YFDMAEDTIIFHM-MGEPHETRLVLMHNEQAVISPSWS---IHT 246

QY 327 GMEHKNLTMVPRINWRIFKEGSGSQPDNDPDLIDYCYGDVR 368
   :||: :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 247 VGVTKNYAF---IWGMTI---GENLTFDD---MDHIAMLDLR 278

RESULT 12
000443 PRELIMINARY; PRT; 1686 AA.
ID 000443
AC 000443
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE PHOSPHOINOSITIDE 3-KINASE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97479209; PubMed=9337861;
RA Domin J., Pages F., Volinia S., Rittenhouse S.E., Zvelebil M.J.,
RA Stein R.C., Waterfield M.D.;
RT "Cloning of a human phosphoinositide 3-kinase with a C2 domain that
RT displays reduced sensitivity to the inhibitor wortmannin.";
RL Blochem. J. 326:139-147(1997).
DR EMBL: Y13367; CAA73797.1; -.
DR HSP; P21707; IBYN.
DR InterPro: IPR000008; C2.
DR InterPro: IPR001263; PI3Ka.
DR InterPro: IPR002420; PI3K_C2.
DR InterPro: IPR000341; PI3K_ras_bind.
DR InterPro: IPR000403; PI3_P14_kinase.
DR InterPro: IPR001683; PX.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00613; PI3Ka; 1.
DR Pfam; PF00792; PI3K_C2; 1.
DR Pfam; PF00794; PI3K_rbd; 1.
DR Pfam; PF00454; PI3_P14_kinase; 1.
DR Pfam; PF00787; PX; 1.
DR SMART; SM00239; C2; 2.
DR SMART; SM00145; PI3Ka; 1.
DR SMART; SM00146; PI3K; 1.
DR SMART; SM00142; PI3K_C2; 1.
DR SMART; SM00144; PI3K_rbd; 1.
DR SMART; SM00312; PX; 1.
DR PROSITE; PSS0004; C2_DOMAIN_2; 1.
DR PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR PROSITE; PS00916; PI3_4_KINASE_2; UNKNOWN_1.
DR PROSITE; PSS0290; PI3_4_KINASE_3; 1.
SQ SEQUENCE 1686 AA; 190736 MW; E9311C803025C96F CRC64;
```

```
Query Match 5.0%; Score 117; DB 4; Length 1686;
Best Local Similarity 19.2%; Pred. No. 0.48;
Matches 86; Conservative 67; Mismatches 154; Indels 140; Gaps 21;

QY 45 QDNELPIDV-----QSATQASASTDTANPLDE-----HEPELYTTALENKTMLIN----- 88
   :||: :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 522 EDEETPVDLNKHLYQIEKPKCEAMTRHPVEELDSYHNQVELALQIENOHRAVDQVIKAV 581

QY 89 --CSALNQDIMRLACYDTLVHGETPAVIKTRSRRLDETIWQTIKKGQVYIYQETDPI 145
   :||: :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 582 RKICSAID-GVETLAI-----TESVKKLRAVNLPSS-----KTADVT 618

QY 146 FLMGNEKGMILTKKDAKQLEYAAKQFTPLSLFDLDRNNTPLWSSRPHNPMYV-----LP 199
   :||: :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 619 SLFGGE-----DPSRST-RGSLNPNPQVQSVINQLTAA 651

QY 200 IF-----MHGPNRSPNTPSHEAKQFTPNFRAPELKFOVSVKVKAEDL---WGTDSDLW 252
   :||: :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 652 IYDLLRLHANSGRSPDCAQSSKSVKEAWTTTEOLOFTIF-----AAHGISSNMYSYNEKY 707

QY 253 FGYYTQOQSHWOIFNGKN-SRPF---RVHDYOPEIFLQO-----PV-YSDLPWGKVRMI 300
   :||: :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 708 YLCISLH-----NGKDLFKPIQSKKVGTYRNFYLIKWDELIIPIQISLPLESVLHLT 763

QY 301 GMGAVRHSHNGESAKLSRSWN-----RAYLMAGMEWKNLTYMPRIWGRIFK 345
   :||: :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 764 LFGILNQSSGSSPDNSKQKRGPEALGVSLPLCDFRRLTCG-----TKLLYLW----- 812

QY 346 EGSQSQPDNDPDLIDYCYGDVRELYOLENKSNTSGVRYNPRSGKALQLDYIYPIGKG 405
   :||: :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 813 --TSSHTNSVPGIVTKKGYMYERIVLQVDFPSPAFDIITPQVDRSIIQHNLETEND 870

QY 406 ISGVYFQIFQYGGQSLIDYNHEATSGV 432
   :||: :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 871 IKG-----KLLDLHKDSSLGL 887

RESULT 13
Q63485 PRELIMINARY; PRT; 602 AA.
ID Q63485
AC Q63485;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE C-RAF ACTIVATED ONCOGENE FUSION PROTEIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87172791; PubMed=3550433;
RA Ishikawa F., Takaku F., Nagao M., Sugimura T.;
RT "Rat c-raf oncogene activation by a rearrangement that produces a
RT fused protein.";
RL Mol. Cell. Biol. 7:1226-1232(1987).
CC -1- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: M15428; AAA2002.1; -.
DR InterPro: IPR000719; Euk_kinase.
DR InterPro: IPR002290; Ser_thr_kin_actslte.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00221; STYKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Oncogene; Serine/threonine-protein kinase;
KW Transference.
SQ SEQUENCE 602 AA; 69209 MW; 88B7BFA90FFB02AC CRC64;

Query Match 4.8%; Score 112.5; DB 11; Length 602;
Best Local Similarity 19.3%; Pred. No. 0.27;
Matches 93; Conservative 71; Mismatches 174; Indels 145; Gaps 23;
```

Db	6	TQOLISSILVL-LLSLPTTALAAGNTRDNFKHLLGNDNVKRSEAGALQLEVDQO	64
QY	66	NPL-DEHEPELY-----TTALENKTMLNCSAL-----NODIMLRACY-DTLVHGTEP	111
Db	65	MTLVQDHGEXIQLRGWSTHGLWFPEITLNDNAYKALANDWESNMIRLAMYVGENGYSNP	124
QY	112	AVLTKRSIRLDETI-----WQT-----IKGKPOVLY	138
Db	125	ELKSRVWIKGIDIAIENDMTYVVDVHVHAPDPRDPVYAGAEDFFRDIAALYPNNPHIY	184
QY	139	QETTPD-----IFLMGNEKGLTKKDAKQLYEAAKQFTPLSLSPDLDRN-----NTPLW	187
Db	185	ELANEPSSNNNGGAGIPNNEEGNAVK-----EYADPIVEMLRDGSNADDDNIIVGSPNW	239
QY	188	SSRP-----HNPWVLPPIF--MHGKPNES--PNTPSHEAKQFTNPEFAPELKFQ	233
Db	240	SQRPDLAANDPIDDHTMTYVHFTGSHAASTESYPETPNSERGVNMSTRYA-----LE	295
QY	234	VSVKVAAEDLWGT-----DSDLWFYCTQ-----SHWQIFNGKNS-----RP	271
Db	296	NGVAVFATE--WGTSQLANGDGGPYFDEADWIEFLNENNINSWANSLTN-KNEVSCAFTP	352
QY	272	FRVH-----DYQPEIFLTQPYSDLPW---D	294
Db	353	FELGKSNATSLDFGPQVQVWVPEELSUSGEYVRARIKGVNYP---IDRTKYTKVLDFDND	409
QY	295	GKYVRMIGMGAHVHSGES-----AKLSRSWNRAYLMAGMEKNLTVMPRING	341
Db	410	GTKQGEV-----NGDSPVEDVVIENEGALKLSGLDASNDVSEGNYYANARLSADGWG	463
QY	342	R	342
Db	464	K	464

```

RESULT 15
Q9F216
ID Q9F216 PRELIMINARY; PRRT 824 AA.
AC Q9F216;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)
DE CELLULOSE.
OS Bacillus sp.
OC Bacteria; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1409;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=KSM-S237;
RX MEDLINE=21036886; PubMed=11193393;
RA Hakamada Y., Hatada Y., Koike K., Yoshimatsu T., Kawai S.,
RA Kobayashi T., Ito S.;
RT "deduced amino acid sequence and possible catalytic residues of a
RT thermostable, alkaline cellulase from an alkaliophilic Bacillus
RT strain.";
RL Btosc1. Biotechnol. Biochem. 64:2281-2289(2000).
RL ENBL: AB018420; BAB19360.1; -.
DR InterPro: IPR001547; Glyco_hydro_F5.
DR Pfam: PF00150; cellulase; 1.
DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
SQ SEQUENCE 824 AA; 91564 MW; 65FA940FE1D729B9 CRC64;

```

Query Match 4.6%; Score 109; DB 2; Length 824;
Best Local Similarity 20.0%; Pred. No. 0.85;
Matches 96; Conservative 59; Mismatches 159; Indels 166; Gaps 28;

[illegible]

```

Db 66 MTLVDQHGKQLRGMSHGLQFPETLNDNAYKALSNDWDSNMRLAMY-----VGENGX 121
QY 110 --TPAVIKTK-----RSIRLDETI--WQT-----IKGKP 134
Db 122 ATNPELIKQVRIDGIELAIEDMVIIVDWHVHAPGDRDPVYAGAKDFFREIAALYPNP 181
QY 135 QVIYQETDP-----IFLMGNEKMLTKKDAKQLEYAAKQFTPLSLSFDLDRN-----N 183
Db 182 HIIYELANEPSSNNGGAGIPNNEEGKAVK-----EYADPIVEMLRKSGNADDNIIIVG 236
QY 184 TPLWSSRP-----HNPVIYLPFPMHGKPNRSPNTPSHEAKQFTPNFRAPELK-- 231
Db 237 SPNWSQRPDLAADNPIDDHHTMTVHFYGTGSHAATESYPSE-----TPNSERGNVMSNT 291
QY 232 ---FOVSVKVKAAEDLWGT-----DSDLWFGYTQQ-----SHWQIENGKNS-- 269
Db 292 RYALENGVAVFATE--WGTQASGDGPGPYFDEADVWIEFLNENNISWANNSLTN-KNEVS 348
QY 270 ---RPFVR-----HDYQP-EIFLT-----QPV-----YSDLFW-- 293
Db 349 GAFTPFELGKSNATNLDGPDHVAPEELSLSGEYVRARIKGVNYEPIDRTKYTKVLWDF 408
QY 294 -DGKVRMIGCAVHHSNGESAKLSRSNRAYLMAGME-----WKNLTVMPRIWGR 342
Db 409 NDGTKQ--GFGVNSDSPNKELIAVDNENNTLKVSGLDVSNVDSDGNFWANARLSANGWGK 466

```

Search completed: November 30, 2001, 14:26:25
Job time: 562 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:16:58 ; Search time 72.04 Seconds
(without alignments)
454.475 Million cell updates/sec

Title: US-09-787-083-4
Perfect score: 2360
Sequence: 1 MKVSLSTLTSLILPCFAILA.....YNHEATSGVGLMLNDWAGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 522463 seqs, 74073290 residues
Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101.*
1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2360	100.0	442	21	AA1985269 BASB034 amino acid
2	2339	99.1	442	21	AA1985271 BASB034 amino acid
3	2330	98.7	442	21	AA1985268 BASB034 amino acid
4	2330	98.7	442	21	AA1985270 BASB034 amino acid
5	822	34.8	370	21	AA1985156 Neisseria gonorrhoe
6	815	34.5	370	21	AA1985157 Neisseria meningit
7	815	34.5	370	21	AA1985158 Neisseria meningit
8	815	34.5	370	21	AA1985159 Neisseria meningit
9	812	34.4	375	21	AA1985228 Neisseria meningit
10	246.5	10.4	355	19	AA198871 H. pylori GHPO 172
11	239.5	10.1	356	19	AA198960 H. pylori ORF 07ap

12	156.5	6.6	253	18	AAW20760	H. pylori outer me
13	109.5	4.6	824	21	AA23180	Bacillus sp. KSM-S
14	109	4.6	800	8	AA270420	Sequence encoded b
15	109	4.6	822	13	AA28021	Alkaline cellulase
16	107	4.5	157	18	AAW20538	H. pylori outer me
17	106	4.5	798	21	AA240925	Human ORF ORF689
18	106	4.5	798	21	AA190225	Human PGC-1 protei
19	106	4.5	798	22	AA283944	Peroxisome prolif
20	105.5	4.5	537	22	AA291000	C glutamicum prote
21	102.5	4.3	1886	19	AA270991	Human class II p13
22	98.5	4.2	467	20	AA13378	Amino acid sequen
23	98.5	4.2	467	21	AA194860	Human protein clon
24	98.5	4.2	467	22	AA288401	Human membrane or
25	98.5	4.2	467	22	AA280246	Human PRO260 prote
26	98.5	4.2	1726	18	AA238756	Phosphatidyl inosi
27	97	4.1	761	20	AA299084	Non-B, non-C, non-
28	95.5	4.0	502	22	AA290031	C glutamicum prote
29	95.5	4.0	516	22	AA278876	C glutamicum SRT
30	95	4.0	522	22	AA295616	Human protein sequ
31	94	4.0	372	16	AA269607	Gonococcal porin-5
32	93.5	4.0	467	19	AA275057	Human secreted pro
33	92	3.9	888	22	AA270751	S cerevisiae apopt
34	92	3.9	1484	12	AA217749	Human alpha-2 macr
35	91.5	3.9	857	13	AA229814	S receptor kinase
36	91	3.9	600	22	AA275092	Human colon cancer
37	91	3.9	1227	22	AA281501	S. epidermidis ope
38	91	3.9	1474	21	AA297157	Human alpha-2-macr
39	91	3.9	1474	22	AA250673	Human alpha-2 macr
40	90	3.8	522	22	AA295513	Human protein sequ
41	90	3.8	564	21	AA210457	T. matsutake pyran
42	90	3.8	564	21	AA281952	Trichoderma derive
43	90	3.8	790	22	AA293045	Human protein sequ
44	90	3.8	841	20	AA24318	Mouse dephosphoryl
45	90	3.8	1627	16	AA267538	Cytadhesin Pl. My

ALIGNMENTS

RESULT 1
AA1985269
ID AA1985269 standard; Protein; 442 AA.
AC AA1985269;
DT 29-JUN-2000 (first entry)
XX
DE BASB034 amino acid sequence #2.
XX
KW Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW hearing loss; antibacterial drug.
XX
OS Moraxella catarrhalis.
XX
EN WO200015802-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99WO-EP06781.
XX
PR 14-SEP-1998; 98GB-0020002.
XX
PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruelle J;
XX
DR WPI: 2000-271440/23.
DR N-PSDB; AA10701.
XX
PT Novel BASB034 polynucleotides and polypeptides from Moraxella
catarrhalis used to prepare vaccines against bacterial infections

XX Claim 3; Page 67; 106pp; English.
 XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 CC strain Mc2908. The invention relates to BASB034 polypeptides from
 CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
 CC polynucleotides and polypeptides may be employed as research reagents and
 CC material for the discovery of treatments and diagnostics for diseases,
 CC particularly human diseases. They are particularly used to diagnose and
 CC treat M. catarrhalis infections. They can be used for diagnosis of
 CC disease, staging of disease, or determining response of an infectious
 CC organism to drugs. The polynucleotides may be used as a source for
 CC hybridization probes, and for screening of genetic mutations, serotype,
 CC organism or strain identification, identification of mutations in BASB034
 CC sequences, and as components of arrays which are useful for diagnostic
 CC and prognostic purposes. The polypeptides can be used to produce
 CC antibodies. The polypeptides can also be used in vaccine formulations,
 CC and to identify agonists and antagonists. The polypeptides, antibodies,
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.
 XX SQ Sequence 442 AA;

Query Match 100.0%; Score 2360; DB 21; Length 442;
 Best Local Similarity 100.0%; Pred. No. 2.8e-219;
 Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKVSLSTLTSLPCFALLAIQQAQVNPVAFVDEVRSKNDLQGNELLIGVQSATQSA 60
 DB 1 mkvslstltslpcfallaiqqaqvnpvafvdevrskndlqgdnelligvqsatqsa 60
 QY 61 STDTPANPLDEHEPELYTTALENKTMLNCSALNODIMRLACYDPLVHCETPAVTKTRSI 120
 DB 61 stdtanpldehepelyttalenktmlncsalnqdimrlacydplvhetpavtktrsi 120
 QY 121 RLDETWTQIKGKQVYQETTPDIFLMGNEKGMMLTKKDAKQLEYAARQFTPLSLSFOLD 180
 DB 121 rldetiwtikgkqvvyqettpdiflmgnekgmmltkkdakqleyaakqftplsfsld 180
 QY 181 RNNVPLSSRPHNPMYVLPIMHGKPNRSPNTPSHEARQFTPNFRAPELKFQVSVKVA 240
 DB 181 rnnvplssrphnmyvlpimhkgpnrsptpshearqftpnfrapelekfvsvkva 240
 QY 241 AEDLWGTSDSLWFGVYQSQHWQIFNGKNSRPFVRVHDYQPEIFLQPVYSDLPWQKVRMI 300
 DB 241 aedlwgtstdlwfgytqsqhwqifngknsrpfvrhdyqpeiflcqpvysdlpwqgkvrmi 300
 QY 301 GMGAVHSHNGESAKLSRSWNRAYLMAGMEWKNLTVMPRINGRIEKGSGSQDPDNPDLID 360
 DB 301 gmgavhshngesaklsrswnraylmagmewknlvtmpringrifkegsgsqdpdpndpild 360
 QY 361 YYGVDVRFYQLENKSNISTVRYNPRSGKALQLODYVPLGLGIGSYFOIFOGYQGS 420
 DB 361 yygvdvrflyqlenksnigtvrynprsgkalgldyvvyplgkigsgyfgifggygqsl 420
 QY 421 IDYNHEATSFVGLMLNDWMGL 442
 DB 421 idynheatsfvglmlndwmgl 442

RESULT 2
 AAY85271
 ID AAY85271 standard; Protein; 442 AA.
 XX
 AC AAY85271;
 XX

DT 29-JUN-2000 (first entry)
 XX BASB034 amino acid sequence #4.
 DE
 XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;
 KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KW hearing loss; antibacterial drug.
 XX
 OS Moraxella catarrhalis.
 XX WO200015802-A1.
 PN 23-MAR-2000.
 PD
 XX 14-SEP-1999; 99WO-EP06781.
 PF
 XX 14-SEP-1998; 98GB-0020002.
 PR
 XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 PA
 XX Ruelle J;
 PI
 XX WPI; 2000-271440/23.
 DR N-PSDB; AAA10703.
 DR
 XX Novel BASB034 polynucleotides and polypeptides from Moraxella
 PT catarrhalis used to prepare vaccines against bacterial infections
 PT
 XX Claim 3; Page 69; 106pp; English.
 PS
 XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 CC strain Mc2969. The invention relates to BASB034 polypeptides from
 CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
 CC polynucleotides and polypeptides may be employed as research reagents and
 CC material for the discovery of treatments and diagnostics for diseases,
 CC particularly human diseases. They are particularly used to diagnose and
 CC treat M. catarrhalis infections. They can be used for diagnosis of
 CC disease, staging of disease, or determining response of an infectious
 CC organism to drugs. The polynucleotides may be used as a source for
 CC hybridization probes, and for screening of genetic mutations, serotype,
 CC organism or strain identification, identification of mutations in BASB034
 CC sequences, and as components of arrays which are useful for diagnostic
 CC and prognostic purposes. The polypeptides can be used to produce
 CC antibodies. The polypeptides can also be used in vaccine formulations,
 CC and to identify agonists and antagonists. The polypeptides, antibodies,
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.
 XX SQ Sequence 442 AA;

Query Match 99.1%; Score 2339; DB 21; Length 442;
 Best Local Similarity 99.1%; Pred. No. 3e-217;
 Matches 438; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MKVSLSTLTSLPCFALLAIQQAQVNPVAFVDEVRSKNDLQGNELLIGVQSATQSA 60
 DB 1 mkvslstltslpcfallaiqqaqvnpvafvdevrskndlqgdnelligvqsatqsa 60
 QY 61 STDTPANPLDEHEPELYTTALENKTMLNCSALNODIMRLACYDPLVHCETPAVTKTRSI 120
 DB 61 stdtanpldehepelyttalenktmlncsalnqdimrlacydplvhetpavtktrsi 120
 QY 121 RLDETWTQIKGKQVYQETTPDIFLMGNEKGMMLTKKDAKQLEYAARQFTPLSLSFOLD 180
 DB 121 rldetiwtikgkqvvyqettpdiflmgnekgmmltkkdakqleyaakqftplsfsld 180

QY 181 RNTPLWSSRPHNPMYVLPFIHMGKPNRSPNTPSHEARQFTPNEFRAPELKFQVSVKVA 240
 DB 181 rntplwssrphnmpvlpfihmgkpnrsntpsheakftnefrapelkfqsvsvkva 240
 QY 241 AEDLWGTDSLDLWFGYTOQSHWQIFNGKNSRPRVRHDYQPEIFLTQPVYSDLPWDGKVRMI 300
 DB 241 aedlwgtddslwfgytqgshwqifngknsrprvrhdypqeilftqpvysdlpwdgkvrmi 300
 QY 301 GMGAVHSHNGESAKLSRSWNRAYLMAGMEWKNLTVMPTWGRIFKFGSGSQPDNDPDIID 360
 DB 301 gmgavhshngesaksrswnraylmagmewknlvtmptwgrifkfgsgsqpdpndpdiid 360
 QY 361 YYGYGDRFLYOLENKSNSISGTVRYNPRSGKALQLDYVYPLGKISGYSYFQIFGQYGSLS 420
 DB 361 yygygdvrfllyqlenksnsgtvyrynprsgkalgldyvypplgkigsgyfifqgygqsl 420
 QY 421 IDYNHEATSFYGLMLNDWMGL 442
 DB 421 idynheatsfgvglmndwmg1 442

RESULT 3
 AAY85268 standard; Protein; 442 AA.

AC AAY85268;

DT 29-JUN-2000 (first entry)

XX BASB034 amino acid sequence #1.

DE Moraxella catarrhalis infection; BASB034; diagnosis; staging;
 KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KW hearing loss; antibacterial drug.

XX Moraxella catarrhalis.

XX WO200015802-A1.

XX 23-MAR-2000.

XX 14-SEP-1999; 99WO-EP06781.

XX 14-SEP-1998; 98GB-0020002.

XX (SMTK) SMITHKLINE BEECHAM BIOLOGICALS.

PI Ruelle J;

DR WPI; 2000-271440/23.

DR N-PSDB; AAA10700.

PT Novel BASB034 polynucleotides and polypeptides from Moraxella

PS catarrhalis used to prepare vaccines against bacterial infections

PS Claim 3; Fig 2; 106pp; English.

XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 CC strain Mc2931 (ATCC 43617). The invention relates to BASB034 polypeptides
 CC from M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The
 CC BASB034 polynucleotides and polypeptides may be employed as research
 CC reagents and material for the discovery of treatments and diagnostics for
 CC diseases, particularly human diseases. They are particularly used to
 CC diagnose and treat M. catarrhalis infections. They can be used for
 CC diagnosis of disease, staging of disease, or determining response of an
 CC infectious organism to drugs. The polynucleotides may be used as a source
 CC for hybridization probes, and for screening of genetic mutations,
 CC serotype, organism or strain identification, identification of mutations
 CC in BASB034 sequences, and as components of arrays which are useful for
 CC diagnostic and prognostic purposes. The polypeptides can be used to
 CC produce antibodies. The polypeptides can also be used in vaccine

CC formulations, and to identify agonists and antagonists. The polypeptides,
 CC antibodies, agonists and antagonists (which are bacteriostatic) are used
 CC for the treatment and prevention of diseases such as otitis media in
 CC infants and children, pneumonia in elderly, sinusitis, nosocomial
 CC infections and invasive diseases, and chronic otitis media with hearing
 CC loss. The polypeptides, agonists and antagonists are also used for
 CC screening of antibacterial drugs. The BASB034 products of the invention
 CC can be used screen for new antibacterial compounds that may target
 CC resistant bacteria.
 XX
 SQ Sequence 442 AA;

Query Match 98.7%; Score 2330; DB 21; Length 442;
 Best Local Similarity 98.6%; Pred. No. 2.2e-216;
 Matches 436; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKVSLTLTSLILPCFALLAIQQAQAVPNPVAFAVDEVRKNDLQDNELLIGVOSTQSA 60
 DB 1 mkvsltltlslilscfaillaiqgaqavpnpvafvdevrsendlqgdneipdvqsatqsa 60

QY 61 STDTANPLDEHEPELYTTALENKTMLNCSALNQDITMLRACYDTLVHGETPAVIKTRSI 120
 DB 61 stdtanpldehepelyttalenktmlnscalnqdmrlacydtlvhgetpaviktkrsi 120

QY 121 RLDETIWQTIKGPVVOYQETTPFIPLMGNEKGMKTKDAKOLEYAAKQFTPLSLSFOLD 180
 DB 121 rldetiwtikgkpvvyoyettppifplmgnekgmktkdaqlkleyaakqftplslsfold 180

QY 181 RNTPLWSSRPHNPMYVLPFIHMGKPNRSPNTPSHEARQFTPNEFRAPELKFQVSVKVA 240
 DB 181 rntplwssrphnmpvlpfihmgkpnrsntpsheakftnefrapelkfqsvsvkva 240

QY 241 AEDLWGTDSLDLWFGYTOQSHWQIFNGKNSRPRVRHDYQPEIFLTQPVYSDLPWDGKVRMI 300
 DB 241 aedlwgtddslwfgytqgshwqifngknsrprvrhdypqeilftqpvysdlpwdgkvrmi 300

QY 301 GMGAVHSHNGESAKLSRSWNRAYLMAGMEWKNLTVMPTWGRIFKFGSGSQPDNDPDIID 360
 DB 301 gmgavhshngesaksrswnraylmagmewknlvtmptwgrifkfgsgsqpdpndpdiid 360

QY 361 YYGYGDRFLYOLENKSNSISGTVRYNPRSGKALQLDYVYPLGKISGYSYFQIFGQYGSLS 420
 DB 361 yygygdvrfllyqlenksnsgtvyrynprsgkalgldyvypplgkigsgyfifqgygqsl 420

QY 421 IDYNHEATSFYGLMLNDWMGL 442
 DB 421 idynheatsfgvglmndwmg1 442

RESULT 4
 AAY85270 standard; Protein; 442 AA.

AC AAY85270;

DT 29-JUN-2000 (first entry)

XX BASB034 amino acid sequence #3.

XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;
 KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KW hearing loss; antibacterial drug.

XX Moraxella catarrhalis.

XX WO200015802-A1.

XX 23-MAR-2000.

XX 14-SEP-1999; 99WO-EP06781.

```

PR      14-SEP-1998;          98GB-00200002.
XX
XX      (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX      Ruelle J;
XX
XX      WPI: 2000-271440/23.
DR      N-PSDB; AAAI0702.
XX
XX      Novel BASB034 polynucleotides and polypeptides from Moraxella
PT      catarrhalis used to prepare vaccines against bacterial infections -
XX
XX      Claim 3; Page 68; 106pp; English.
XX
XX      This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC      strain Mc2913. The invention relates to BASB034 polypeptides from
CC      M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2959. The BASB034
CC      polynucleotides and polypeptides may be employed as research reagents and
CC      material for the discovery of treatments and diagnostics for diseases,
CC      particularly human diseases. They are particularly used to diagnose and
CC      treat M. catarrhalis infections. They can be used for diagnosis of
CC      disease, staging of disease, or determining response of an infectious
CC      organism to drugs. The polynucleotides may be used as a source for
CC      hybridization probes, and for screening of genetic mutations, serotype,
CC      organism or strain identification, identification of mutations in BASB034
CC      sequences, and as components of arrays which are useful for diagnostic
CC      and prognostic purposes. The polypeptides can be used to produce
CC      antibodies. The polypeptides can also be used in vaccine formulations,
CC      and to identify agonists and antagonists. The polypeptides, antibodies,
CC      agonists and antagonists (which are bacteriostatic) are used for the
CC      treatment and prevention of diseases such as otitis media in infants and
CC      children, pneumonia in elderly, sinusitis, nosocomial infections and
CC      invasive diseases, and chronic otitis media with hearing loss. The
CC      polypeptides, agonists and antagonists are also used for screening of
CC      antibacterial drugs. The BASB034 products of the invention can be used
CC      screen for new antibacterial compounds that may target resistant
XX      bacteria..
XX
XX      Query           442 AA;
SQ
Query Match              98.7%; Score 2330; DB 21; Length 442;
Best Local Similarity   98.9%; Pred. No. 2.2e-216;
Matches 437; Conservative    2; Mismatches 3; Indels 0; Gaps 0
QY      1 MKVSLTSLTLSPCFAILAQOAAVPNFAFDEVRSRNDLGQDRELLIGVSATQSA 60
Db      1 mkvsltltsllscfaillaiqakavnpfafdevrsendlgqdnlpldvqsatgsa 60
QY      61 STDTANPLDEHEPELYTTALENTKMLINCSALNQDINRLACYDTLVHGEPPAVIKTKRSI 120
Db      61 stdtanpldehepeyltalenktmlincsalngdlimrlacydtlvhgepaviktkrsi 120
QY      121 RLDETWTQTIKGPQVVYQETTDPIFLMGNEKGMLTKDKAKOLEYAARKQTFTPLSLSDLD 180
Db      121 rldetiwtqtkgpqvvygettddpiflmgnekgmltkdkakoleyaarkqtftplsldld 180
QY      181 RNNTPLWSSRRPHNPMTYLPFIWMHGKNRSPNTPSSEARQFTPNFEPAELKFQVSVKVKA 240
Db      181 rnntplwssrrphnmpyvlpfimhngkpnxspntpsshearqfcpnfeprapelkfqvsvkvka 240
QY      241 AEDLWGTDSDLWFGYTQQSHQIFNGKNSRFVRHDYQPETFLTPQVYSDLPWDGKVRMI 300
Db      241 aedlwgtddslwfgytcqshwqlfngksrpfvrhdypetfltpqvysdlpwdgkvrmi 300
QY      301 GMGAVHHSNGESAKLSRSWNRAYLMAGMEWKNLTVMPRIWGRIFKEGGSGOPDDNPDIID 360
Db      301 gmgavhhsgesakerswnraylmagmewknlvmprirwgrifkeggsgopddnpiid 360
QY      361 YYGYGDVRFYLQLENKSNISGTFRYNRPSSRGKALQLDYVYPLGKIGSYFOIFQYGQSIL 420
Db      361 yygygdvrflyqlenksnlsgtvrynpssrgkalqldyvyplogkigsyfoifqygqsil 420

```

QY	421	IDYNHEATSFQVGLMLNDWMGL 442
Db	421	idynheatsfgvglmndwmg1 442
RESULT	5	
AAV75156		
ID	AAV75156 standard; Protein; 370 AA.	
XX	AC	AAV75156;
XX	DT	21-MAR-2000 (first entry)
XX		Neisseria gonorrhoeae ORF 582 protein sequence SEQ ID NO:1786.
XX		Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW		antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
KW		antibacterial; gene therapy.
XX		
OS		Neisseria gonorrhoeae.
XX		
PN		WO9957280-A2.
XX		
PD		11-NOV-1999.
XX		
PF		30-APR-1999; 99WO-US09346.
XX		
PR		01-MAY-1998; 98US-0083758.
PR		31-JUL-1998; 98US-0094869.
PR		02-SEP-1998; 98US-0098994.
PR		02-SEP-1998; 98US-0099062.
PR		09-OCT-1998; 98US-0103749.
PR		09-OCT-1998; 98US-0103794.
PR		03-OCT-1998; 98US-0103796.
PR		23-FEB-1999; 99US-0121528.
XX		
PA		(CHIR) CHIRON CORP.
PA		(GENO-) INST GENOMIC RES.
XX		
PI		Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI		Petersen J, Pizzo M, Rappuoli R, Ratti G, Scialato E, Scarselli M;
PI		Tettelin H, Venter JC;
XX		
DR		WPI; 2000-062150/05.
DR		N-PSDB; AA253918.
XX		
PT		Novel Neisserial polypeptides predicted to be useful antigens for
PT		vaccines and diagnostics -
XX		
PS		Claim 2; Page 903; 1453pp; English.
XX		
CC		AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAY75941
CC		represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
CC		and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC		PCR primers used in the exemplification of the present invention. The
CC		polypeptides, the polynucleotides, antibodies and compositions of
CC		the invention can be used as vaccines, as diagnostic reagents, and as
CC		immunogenic compositions. The polypeptides can be used in the
CC		manufacture of medicaments for treating or preventing infection due to
CC		Neisserial bacteria (e.g. meningitis and septicemia), to detect the
CC		presence of Neisseria bacteria, or to raise antibodies. They may also
CC		be used to screen for agonists or antagonists, which may themselves
CC		have use as antibacterial agents. The polynucleotides of the invention
CC		may also be used in gene therapy protocols.
XX		
SQ		Sequence 370 AA;
Query Match 34.8%; Score 822; DB 21; Length 370;		
Best Local Similarity 44.7%; Pred. No. 8.5e-71;		
Matches 163; Conservative 68; Mismatches 110; Indels 24; Gaps		
QY	87	INCSALNODIWRACYDPLVHGTEPAVI-----KTKRSIRLDETIVQTI-KGKPPQVVYQE 140

Db	193	sdwlyngqrksapfrntdykpeifltqpvkadlpfgrlrmigagfvhdsnggsrpesr	252
Qy	318	SNRAYLMAGWENKLTVMPIRWGRIFKEGSGQDPDNDPILDYGYGDVRFYQLENKS	377
		: : : :	
Db	253	swriyamagmegkltvprvvrfdq-sgdk-nanpdiaymgvdvklqyrlndrq	310
Qy	378	NISGTVYRPSRGKALQLDYYVPLGKGISGYFOIFOGYGOSLIDYNHEATSFQVGLMLN	437
		: : : : : :	
Db	311	nvysvlrnpktygaaleaaytfpikgklgvvrgfhgygesliidynhkgngigimfn	370
Qy	438	DWVGL 442	
		:	
Db	371	dlldgi 375	
RESULT	10		
AAW98871			
ID	AAW98871	standard; Protein; 355 AA.	
XX	AC	AAW98871;	
XX	DT	31-MAR-1999 (first entry)	
XX	DE	H. pylori GHPO 1723 protein.	
XX	DE	GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;	
KW	KW	peptic ulcer disease.	
XX	XX	Helicobacter pylori.	
OS	XX	WO9843478-A1.	
PN	XX	08-OCT-1998.	
PD	XX	01-APR-1998; 98WO-US06371.	
PF	XX	29-JUL-1997; 97US-0902615.	
XX	PR	01-APR-1997; 97US-0833457.	
XX	PR	24-JUN-1997; 97US-0881227.	
XX	XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	PA	(INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.	
XX	XX	Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;	
PI	XX	WPI: 1998-542293/46.	
XX	DR	N-PSDB; AAX14590.	
XX	XX	New isolated Helicobacter polynucleotides - used to develop products	
PT	PT	for the diagnosis, prevention and treatment of Helicobacter	
PT	PT	infections and gastrointestinal diseases	
XX	XX	Claim 8; Page 1976-1977; 2054pp; English.	
XX	XX	This sequence represents a Helicobacter pylori GHPO protein of the	
CC	CC	invention. The polypeptides can be used for preventing or treating	
CC	CC	Helicobacter infections, and gastroduodenal diseases associated with	
CC	CC	these infections, including acute, chronic, and atrophic gastritis, and	
CC	CC	peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be	
CC	CC	used for the production of antibodies. The products can also be used for	
CC	CC	detection and diagnosis.	
XX	XX	Sequence 355 AA;	
SQ			
Query Match		10.4%; Score 246.5; DB 19; Length 355;	
Best Local Similarity		25.1%; Pred. NO. 3.2e-15;	
Matches		82; Conservative 41; Mismatches 117; Indels 87; Gaps 11;	
Qy	157	KKDAKQLEYAAKQFTPLSLFSFDLDRNNTPLWSSRPHNPMYVLPFIMHCKENRSPNTPSHE	216
		:	
Db	69	kkylnmndyigtvlfpyhsf-----tpifqwyhpninp-----	102


```

DR WPI; 1987-325245/46.
XX N-PSDB; AAN70651.
XX DNA sequence coding cellulase gene - is derived from Bacillus sp. no.
PT 11392 and can hydrolyse cello-triose and cello-tetraose.
XX PS Disclosure; Fig 3-1 - 3-3; 1lpp; Japanese.
XX This cellulase hydrolyses cellotriose or cellotetraose and does not
CC hydrolyse cellobiose. Molecular weight is approx. 92K dalton and
CC optimum pH is 9.
XX Sequence 800 AA;
SQ
Query Match 4.6%; Score 109; DB 8; Length 800;
Best Local Similarity 19.7%; Pred. No. 0.23;
Matches 91; Conservative 51; Mismatches 150; Indels 170; Gaps 24;
QY 30 PVAFVDEVRKND-----LGQDNELLIGVQSATQSASTDTANPL-DEHEPELY-----TT 78
DB 25 ptalaaegntrednfkhllgndnvkrpseagalqlqevdgqmtlvdqhgkqlrgmsth 84
QY 79 ALENKTMNLINCSAL-----NODIMRLACY-DTLVHGTEPAVTKRSIRLDETI----- 126
DB 85 glqwfpeilndnaykalandwesnmirlamyvgengyasnpeliksrvikgidlalendm 144
QY 127 -----WQT-----IKGPKQVYVYQETDP-----IFLMGN 150
DB 145 yvldvwhvhapgdprdpvyagaedffrdiaalypnnpnhiyelaneppsnnggagipnn 204
QY 151 EKGMLTKKDAKQLEYAAKQTPLSLSFDLDRN-----NTPLWSSRP-----HNPMY 196
DB 205 eegwnavk-----eyadpivemirdsgnaddnliivgsnwsgrpdlaadnpiddhhtmy 259
QY 197 VLPIF--MHCKPNRS--PNTPSHEARQFTNEFRAPELKFQVSVKVAADLWGT----- 247
DB 260 tvhfytgshaastesyppetcpnsergnvmsntrya----lengvavfate--wgtsgang 313
QY 248 -----DSDLWFGYTOQ-----SHWQIFNGKNS-----RPFVRH----- 275
DB 314 dgapyfdeadvwieflnenniswanwsltn-knevsagftpfelgksnatsldpgpdqvw 372
QY 276 -----DYQPEIFLTQPVYSDLPW---DGKVRMIGMGAVHHSNGES- 312
DB 373 vpeelslsgeyvrarikgvnyep---ldrtkytkvlwdfndgtkqgfgv-----ngdsp 423
QY 313 -----AKLSRSWNRAYLMAGMEWKNLTVMPRINGR 342
DB 424 vedvvieneagalksgldasndvsegnywanarlsadgwgk 465
RESULT 15
AAR26021
ID AAR26021 standard; Protein; 822 AA.
XX AC AAR26021;
XX DT 02-FEB-1993 (first entry)
XX DE Alkaline cellulase K-64.
XX KW Mass production; recombinant.
XX OS Bacillus sp. KSM-64 (FERM P-10482).
XX PN JP04190793-A.
XX PD 09-JUL-1992.
XX PF 26-NOV-1990; 90JP-0324401.
XX PR 26-NOV-1990; 90JP-0324401.

```

```

XX (KAOS ) KAO CORP.
XX WPI; 1992-280112/34.
XX N-PSDB; AAQ27180.
XX Alkaline cellulase mass prodn. - using recombinant plasmid and
PT microorganism contg. cellulase gene
XX Disclosure; Fig 4; 14pp; Japanese.
XX Alkaline cellulase K-64 may be mass produced recombinantly in host
CC microorganisms such as E.coli HB101, C600 or JM109, B.subtilis B0170,
CC 168 or ISW1214.
XX Sequence 822 AA;
SQ
Query Match 4.6%; Score 109; DB 13; Length 822;
Best Local Similarity 19.7%; Pred. No. 0.24;
Matches 91; Conservative 51; Mismatches 150; Indels 170; Gaps 24;
QY 30 PVAFVDEVRKND-----LGQDNELLIGVQSATQSASTDTANPL-DEHEPELY-----TT 78
DB 25 ptalaaegntrednfkhllgndnvkrpseagalqlqevdgqmtlvdqhgkqlrgmsth 84
QY 79 ALENKTMNLINCSAL-----NODIMRLACY-DTLVHGTEPAVTKRSIRLDETI----- 126
DB 85 glqwfpeilndnaykalandwesnmirlamyvgengyasnpeliksrvikgidlalendm 144
QY 127 -----WQT-----IKGPKQVYVYQETDP-----IFLMGN 150
DB 145 yvldvwhvhapgdprdpvyagaedffrdiaalypnnpnhiyelaneppsnnggagipnn 204
QY 151 EKGMLTKKDAKQLEYAAKQTPLSLSFDLDRN-----NTPLWSSRP-----HNPMY 196
DB 205 eegwnavk-----eyadpivemirdsgnaddnliivgsnwsgrpdlaadnpiddhhtmy 259
QY 197 VLPIF--MHCKPNRS--PNTPSHEARQFTNEFRAPELKFQVSVKVAADLWGT----- 247
DB 260 tvhfytgshaastesyppetcpnsergnvmsntrya----lengvavfate--wgtsgang 313
QY 248 -----DSDLWFGYTOQ-----SHWQIFNGKNS-----RPFVRH----- 275
DB 314 dgapyfdeadvwieflnenniswanwsltn-knevsagftpfelgksnatsldpgpdqvw 372
QY 276 -----DYQPEIFLTQPVYSDLPW---DGKVRMIGMGAVHHSNGES- 312
DB 373 vpeelslsgeyvrarikgvnyep---ldrtkytkvlwdfndgtkqgfgv-----ngdsp 423
QY 313 -----AKLSRSWNRAYLMAGMEWKNLTVMPRINGR 342
DB 424 vedvvieneagalksgldasndvsegnywanarlsadgwgk 465

```

Search completed: November 30, 2001, 14:16:59
Job time: 381 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:17:42 ; Search time 36.79 seconds
(without alignments)
270.358 Million cell updates/sec

Title: US-09-787-083-4
Perfect score: 2360
Sequence: 1 MKVSLSTLTLILPCFAILA.....YNHEATSEFGVGLMLNDWMGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pap.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pap.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pap.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pap.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pap.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	101.5	4.3	1726	2	US-08-609-049A-30
2	101.5	4.3	1726	4	US-09-170-996-30
3	99.5	4.2	1658	2	US-08-609-049A-13
4	99.5	4.2	1658	4	US-09-170-996-13
5	89.5	3.8	857	1	US-07-717-331F-2
6	87.5	3.7	844	1	US-07-731-157A-6
7	87.5	3.7	844	2	US-08-541-780-6
8	87	3.7	355	1	US-07-946-497-5
9	87	3.7	355	2	US-08-483-322-5
10	87	3.7	355	1	US-08-478-882-5
11	87	3.7	503	1	US-07-946-497-2
12	87	3.7	503	1	US-08-483-322-2
13	87	3.7	503	2	US-08-478-882-2
14	86	3.6	1088	3	US-08-633-768A-1
15	84.5	3.6	1024	4	US-09-091-117-5
16	84	3.6	522	6	RE34506-6
17	84	3.6	1627	1	US-07-665-792E-9
18	83	3.5	494	1	US-08-275-488A-4
19	83	3.5	494	1	US-08-275-490-4
20	83	3.5	494	1	US-08-446-380-4
21	83	3.5	494	1	US-08-446-374-4
22	83	3.5	494	1	US-08-446-382-4
23	83	3.5	494	1	US-08-445-801-4
24	83	3.5	494	1	US-08-275-487-4
25	83	3.5	494	5	PCT-US95-08919-4
26	83	3.5	774	3	US-08-902-632-2
27	83	3.5	774	3	US-09-073-354-1

28	83	3.5	774	3	US-08-656-005A-1	Sequence 1, Appl
29	83	3.5	774	4	US-09-073-259-1	Sequence 1, Appl
30	83	3.5	774	4	US-09-363-095-1	Sequence 1, Appl
31	83	3.5	774	4	US-09-418-027-1	Sequence 1, Appl
32	82.5	3.5	657	4	US-09-306-593-2	Sequence 2, Appl
33	82	3.5	537	2	US-08-633-879C-2	Sequence 2, Appl
34	82	3.5	770	1	US-08-445-135-2	Sequence 2, Appl
35	81.5	3.5	663	1	US-08-441-139-7	Sequence 7, Appl
36	81.5	3.5	1022	3	US-08-772-270A-2	Sequence 2, Appl
37	81	3.4	649	2	US-08-871-266B-16	Sequence 16, Appl
38	81	3.4	649	2	US-09-018-864A-16	Sequence 16, Appl
39	81	3.4	649	3	US-08-871-267B-22	Sequence 22, Appl
40	81	3.4	649	4	US-09-618-419-22	Sequence 22, Appl
41	81	3.4	1121	1	US-07-789-915A-2	Sequence 2, Appl
42	81	3.4	1121	1	US-08-005-002C-2	Sequence 2, Appl
43	81	3.4	1121	1	US-08-487-203A-2	Sequence 2, Appl
44	80	3.4	334	2	US-08-359-850-4	Sequence 4, Appl
45	80	3.4	438	1	US-07-923-095-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-609-049A-30
; Sequence 30, Application US/08609049A
; Patent No. 5948664
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Molz, Lisa
; APPLICANT: Chen, Yen-Wen
; TITLE OF INVENTION: No. 5948664e1 PI 3-Kinase Polypeptides
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,049A
; FILING DATE: 29-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-063700US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1726 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-609-049A-30

Query Match 4.3%; Score 101.5; DB 2; Length 1726;
Best Local Similarity 18.6%; Pred. No. 0.43;
Matches 98; Conservative 79; Mismatches 190; Indels 161; Gaps 21;
QY 1 MKVSLS-----TLTSLILPCFAILAIOQAQVNPVAFVD-----EVRSKNDLQGD 46
DB 465 KVKVSTIEGLQLPVTFTCDVSVSTVEIIMQALSWSHDDLNOVDVGSYILKVCQGEVLQN 524


```

; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/170,996
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/609,049
; FILING DATE: 29-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-063700US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1658 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-170-996-13

Query Match 4.2%; Score 99.5; DB 4; Length 1658;
Best Local Similarity 18.6%; Pred. No. 0.65;
Matches 98; Conservative 79; Mismatches 190; Indels 161; Gaps

QY 1 MKVSL-----TLTSLPCFAILAIQQAQVNPVAFVD-----EVRSKNDLGQD 46
Db 397 KVSIEIEGLQLPVTFCDVSSVTEIIMQALCWVHDDLNVQVDGVSILKVCQEEVLQ 456
QY 47 NELLIGVQSQTASDTANPLDEHEPELVTTALENKTMLINGSALNODIMRLACYDTLV 106
Db 457 NHCLGSHEHTQNCRKWD-----EIKQLLTLSMCQNLAETD-----497
QY 107 HGETPAVIKTKRSIRLDETITQITIKGPQVYQETTDPIF-----146
Db 498 --EAP-----VDLNKLYQIEKPYKEVMIRHPVEELDSYHYQVELALQTEHQHRAV 574
QY 147 -----LMGNEKMLTKDKAQLEYAA-----KQFTPLSLSPDLDRNNTPLWS 198
Db 548 DQVIKAVRKTCALSALDGVTETPSVTEA-VKLIKRAVNLPRNKSADVTSLGSDTRKNTKGS 606
QY 189 SRHPNPMVLP-----IFMHGPNR-----SPNTPSHEARQFTPNFEAPELKFOV 234
Db 607 LNPNPQVQSMHDHTTAIYOLLRLHANSRCSCTGCPGRSNIKEAWTATE-----QL 658
QY 235 SVKYKAAEDL---WGTDSDLWFGYTQSHWOIFNGKN-SRPF---RVHDYQPEIFLTQ-- 285
Db 659 QFTVYAAHGISSNWVSNVEKYVILICSLSH---NGKDLFKPIQSKKVGTVKNEFYLIKWD 714
QY 286 -----PV-YSDLPWDGKVRMIGVAHVHSNGESAKLSRSWN-----RAYL 324
Db 715 ELIIFPIQISQLPLESVLHTLFGLVNLQSSGSSPDSNQRKQGPALGKVSLTLDPKRFL 774
QY 325 MAGMEWKNLNVMPRIWGRIFKEGSGQPDNDPDLIDYGYGDVRFYOLENKNKNISGTVR 384
Db 775 TCG-----TKLLYLW-----TSSHTNSIPGAIPKKSVMERIVLQDFPSPAFDIY 821
QY 385 YNPRSGRGALQDYYVPLRGKISGYFOIFQGYGQSLIDYNHEATSFQV 432
Db 822 TSPQIDRNIIQQDKLETLESIDG-----KLLDIHRDSSFL 859

RESULT 5
US-07-717-331F-2

```

; Sequence 2, Application US/07717331F
; Patent No. 5484905
; GENERAL INFORMATION:
; APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua
; APPLICANT: Stein
; TITLE OF INVENTION: A Receptor Protein Kinase Gene
; TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/717,331F
; FILING DATE: June 19th 1991
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 857 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-717-331F-2

Query Match 3.8%; Score 89.5; DB 1; Length 857;
Best Local Similarity 21.4%; Pred. No. 2.5; Mismatches 110; Indels 119; Gaps 23;
Matches 76; Conservative 50;
QY 105 LVHGETPAVIKTKRSIRLDETITQIKGPKQVVYQETDPIFLMG----NEK---GMLTK 157
Db 24 LIHPALSIVINT-----LSSTESLTSSNKTIV---SPGSIFEVGFTNSRWYLGWYK 75
QY 158 K-DAKQLETAAGKQFTPLSLSD-----LDRNTPPLWSSRPHPNPMYVLPPIPMHCK 205
Db 76 KVSDRTYVYVYANRDNPLSNAIGTLKISGNLVLHDSNKPVMWTN-----LTRGN 125
QY 206 PNRSP-----NTPSHEARQFT-----PNEFRAPELKFQVSVKVAEEDLW 245
Db 126 -ERSPVVAELLANGFNVRDSSNNDAEYLWQSFDPYPTDLPPEMLGYNLKT----- 177
QY 246 GTDSLWFGYTOQSHQIFNGKNSRFRVHYQ-----PEIFLTQ---PVYSDLPWDGK 296
Db 178 GLNREL-----TSRSDDPSSGNF---SYKLEQLSPEFVLSRENPMHRSQPNW- 226
QY 297 VRMIGCAVH-----HSGESAKLSRSWNRAY-----LMAGMEWKNLTVMP--RI 339
Db 227 IRESGIPEDQKLSYMYVNTENNEEVAYFTMTNNSFYSLRLTISEGYFQLTWYPSIRI 286
QY 340 WGRIFEGSGSOPDD---NPDILDYGYCDVRFYOLENKNISGTVR--YNPRS 389
Db 287 WNRFSWSPVDRQCDTYMCGP-----YAYCDV-----NTSPVCNIOGFNPN 329

RESULT 6
US-07-731-157A-6
; Sequence 6, Application US/07731157A
; Patent No. 5457032
; GENERAL INFORMATION:

; APPLICANT: Quax, Wilhelmus J.
; APPLICANT: Misset, Onno
; APPLICANT: Van der Laan, Jan M.
; APPLICANT: Leeting, Herman B.M.
; TITLE OF INVENTION: Mutated beta-lactam acylase genes
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM
; STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/731,157A
; FILING DATE: 19910509
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 90200962
; FILING DATE: 18-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: RAE-VENTER PH.D., BARBARA
; REGISTRATION NUMBER: 32,750
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-494-7622
; TELEFAX: 415-857-0663
; TELEX: 380816 COOLEY PA
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 844 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Kluyvera citrophilia
; STRAIN: ATCC 21285
; US-07-731-157A-6

Query Match 3.7%; Score 87.5; DB 1; Length 844;
Best Local Similarity 21.2%; Pred. No. 3.9; Mismatches 152; Indels 143; Gaps 25;
Matches 92; Conservative 46;
QY 31 VAFVDEVRSK--NDLQD--NEL--LIGVQSATQSASTDTANPLDEHEPELYTAL---- 80
Db 182 LALLTAVKDKYGNDEGMVFNQKLVNPSAPTIAARESSYPLKFDLQNTQTAAALLVPR 241
QY 81 -ENKTMILNCSALNQDIMRLACYDTLVHGETPAVIKTKRSIRLDETITQIKGPKQVVYQ 139
Db 242 YDQAPMLDRPAKGTGALLAV-----TAIKNRETIAA-----QFANGANGLAGY 286
QY 140 ETTDPIFLMGNEKMLTKDAKQLETAAGKQFTPLSLSDLRNTPPLWSSRPHPNPMYVLP 199
Db 287 PTTNSNHWIGKNKA---QDAKAIMVNGPOFG-----W-----YAPATYG 323
QY 200 IFMHGKP-NRSPNTPSHEARQFTNEFRAPELKFQVSVKVAEEDLWGT-----DSDLW 252
Db 324 IGLHAGYDVTGNTP-----FAYPLVFGHNGTIS-----WGSTAGFGDDVDIF 367
QY 253 F-----GYTQOQSHQIFNGKNSRFRVHYQPEIFL-----TQPVYS 289
Db 368 AEKLSAEKPGYYQHNGEWEWVKMLSRKETIAVKDQGETFTVWRTLDGNVKTDRTQTAYA 427
QY 290 DL-PWDGKVRMIGMGAHVHNSGESAKLSRSWNRAYLMAGHEWKNLT-----VMPRIW-- 340
Db 428 KARAWAGK-----EVALS-LAW--THQKAKNWPETWQAAAKQALTINWY 470


```
;
; TYPE: amino acid
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: rat protein
; US-08-483-322-5

Query Match      3.7%; Score 87; DB 1; Length 355;
Best Local Similarity 19.3%; Pred. No. 1.1;
Matches 59; Conservative 36; Mismatches 97; Indels 114; Gaps 16;

QY 45 QNELLIG--VOSATQSASTDTA-NPLDHEPELYTALLENKTMNLNCSALNODIMRLAC 101
   :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
Db 7 RDAFFIGTSLATSTESNTNPTGKPNNEDE-----TDKYPNFGSGIDDD-----ED 56

QY 102 YDTLVHGETPAV-IKTKRSIRLDETIWQIKGPKQVYVQETTDPIFLMGNEKGLTKKDA 160
   :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
Db 57 FSISSATTTPWSAHTKQN--QERTQWNPISNPEVLLQTT-----96

QY 161 KOLEYAAKQFTPLSLSFOLDNRNNTPL-----WSSRPHNPMYVLPFIEMHGKPNRSPNTPSHE 216
   :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
Db 97 -----RMT-----DIDRNSTSAHGENWTQEPQPP-----NNHEYQ 127

QY 217 ARQFTNEFRAPFELKQVSVKVAEDLW-----GTSDLMFGYTOQSHWQIFNG 266
   :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
Db 128 DEEETPH-----ATSTTWADPNSTTTEAAATQEKWF-----ENEMQ--G 164

QY 267 KN-----SRPFRVHDYQPEIFLTQPVYSDLPWDGKVRMIG--MGAVVHHSNGES 312
   :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
Db 165 KNPTPSEDSHVTEGTASAHNNHPSQRMSTQSQEDVSWTDFDPIHPMGQGHQTESKD 224

QY 313 AKLSRS 318
   :|:
Db 225 TGSSES 230

RESULT 10
US-08-483-322-5
; Sequence 5, Application US/08483322
; Patent No. 5760178
; GENERAL INFORMATION:
; APPLICANT: HERRLICH, Peter
; APPLICANT: PONTA, Helmut
; APPLICANT: GUENTHER, Ursula
; APPLICANT: MATZKU, Siegfried
; APPLICANT: WENZL, Achim
; TITLE OF INVENTION: VARIANT CD44 SURFACE PROTEINS, DNA
; TITLE OF INVENTION: SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
; TITLE OF INVENTION: AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,322
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/946,497
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16915/145
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
```

```
;
; TYPE: amino acid
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: rat protein
; US-08-483-322-5

Query Match      3.7%; Score 87; DB 1; Length 355;
Best Local Similarity 19.3%; Pred. No. 1.1;
Matches 59; Conservative 36; Mismatches 97; Indels 114; Gaps 16;

QY 45 QNELLIG--VOSATQSASTDTA-NPLDHEPELYTALLENKTMNLNCSALNODIMRLAC 101
   :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
Db 7 RDAFFIGTSLATSTESNTNPTGKPNNEDE-----TDKYPNFGSGIDDD-----ED 56

QY 102 YDTLVHGETPAV-IKTKRSIRLDETIWQIKGPKQVYVQETTDPIFLMGNEKGLTKKDA 160
   :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
Db 57 FSISSATTTPWSAHTKQN--QERTQWNPISNPEVLLQTT-----96

QY 161 KOLEYAAKQFTPLSLSFOLDNRNNTPL-----WSSRPHNPMYVLPFIEMHGKPNRSPNTPSHE 216
   :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
Db 97 -----RMT-----DIDRNSTSAHGENWTQEPQPP-----NNHEYQ 127

QY 217 ARQFTNEFRAPFELKQVSVKVAEDLW-----GTSDLMFGYTOQSHWQIFNG 266
   :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
Db 128 DEEETPH-----ATSTTWADPNSTTTEAAATQEKWF-----ENEMQ--G 164

QY 267 KN-----SRPFRVHDYQPEIFLTQPVYSDLPWDGKVRMIG--MGAVVHHSNGES 312
   :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
Db 165 KNPTPSEDSHVTEGTASAHNNHPSQRMSTQSQEDVSWTDFDPIHPMGQGHQTESKD 224

QY 313 AKLSRS 318
   :|:
Db 225 TGSSES 230

RESULT 10
US-08-483-322-5
; Sequence 5, Application US/08478882
; Patent No. 585575
; GENERAL INFORMATION:
; APPLICANT: HERRLICH, Peter
; APPLICANT: PONTA, Helmut
; APPLICANT: GUENTHER, Ursula
; APPLICANT: MATZKU, Siegfried
; APPLICANT: WENZL, Achim
; TITLE OF INVENTION: VARIANT CD44 SURFACE PROTEINS, DNA
; TITLE OF INVENTION: SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
; TITLE OF INVENTION: AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,882
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/946,497
; FILING DATE: 19921109
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16915/145
; TELECOMMUNICATION INFORMATION:
```


APPLICATION NUMBER: 9321301.5
FILING DATE: 15-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: DY007.001APC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714-760-0404
TELEFAX: 714-760-9502
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1088 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-633-768A-1

Query Match 3.6%; Score 86; DB 3; Length 1088;
Best Local Similarity 21.6%; Pred. No. 8.5;
Matches 72; Conservative 40; Mismatches 138; Indels 84; Gaps 15;

QY 106 VHGEPAY-IKTKRSIRDEITWOTIKGKPOVYQETDPIFLMGNEKMLTKKDAKOLE 164
DB 29 IHSAPAYATARKTNRLN-VSMTALSDKOTATAGSTD-----NPDGI---DKITYD 76
QY 165 YAAK-QETPLS-----LSFDLRNNTPLWSSRRHNPVYLPIFM 202
DB 77 YGVGAGFSLSTNTWFAAGSSITPGITDWTATMNNFDRIDNP--SITVQHVOQVITY 134
QY 203 HGKPRSPPTSGHEARQPTNEFFRAPELKFQYVSVKAAEDLMGSDIMFGYTOOSHMO 262
DB 135 NNNSYRVEFPNPGPIRDYV---RGPILKQQLD-WIRQELSEGCDDPGM----- 178
QY 263 IFNGKNSRPERHDYQPELFLQPYSDLPMDGKVRMIGMGAHVHNSGSAKLSRWNA 322
DB 179 TPTSGFLTFETKDLVITYIGNFKTRVTRKSDGKVI-----ENDSVTAS----- 224
QY 333 YLMAGEMKNTLVMPRIMGRIKFEKSGSQPDNDPILDYGYGYDFLYO---LENKSN 379
DB 225 ---SGNKRGMLFVDRILGNALIASVKNKFNRAVDVAKQEGFYGAGVCKKQDIILERTOI 281
QY 380 SGT-----VRYN-----PRSGKALQLDYVPL 402
DB 282 AMTNYNYDNLNMQWDLRPPHHDGALNPDYIIPM 315

RESULT 15
US-09-091-117-5
Sequence 5, Application US/09091117
Patent No. 6171589
GENERAL INFORMATION:
APPLICANT: The University of Melbourne
TITLE OF INVENTION: Mycoplasma Recombinant Polypeptides and
TITLE OF INVENTION: Vaccines
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: GREENLEE, WINNER and SULLIVAN P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: United States of America
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/091.117
FILING DATE: 12 JUNE 1998

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU96/00803
FILING DATE: 13-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PN7127
FILING DATE: 13-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: WINNER, Ellen P.
TELECOMMUNICATION INFORMATION:
TELEPHONE: +1 303 499 8080
TELEFAX: +1 303 499 8080
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1024 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Mycoplasma genitalium
US-09-091-117-5

Query Match 3.6%; Score 84.5; DB 4; Length 1024;
Best Local Similarity 20.9%; Pred. No. 11;
Matches 89; Conservative 60; Mismatches 170; Indels 107; Gaps 23;

QY 1 MKVS-LSITLISILPCFAILAIQAAQAVNPVAFVDEVRKNDLGODNELLIGVOSATOS 59
DB 1 MKLSITLISILSICAFGTAL---ALPTVALL-----KHHQOQNTF----- 40
QY 60 ASTDPANPDEHEPELYTALENKTKMLNGSALNDIMRLACYDTLY-HGTPAVIKTR 118
DB 41 ---KOONPTKDIRFGL-----NNVOVPNTIPRHQTVVENVTKNAKAVLDYDKAPQKFLAK 91
QY 119 SIRDEITWOTIKGKPOVYQETDPIFLMGNEKMLTKKDAKOLEYAAKO-FTPLSLF 177
DB 92 S-----ALNKKIQVEFDK-----FL--RTGYINLMDKEMIDQTLFIPQSF 135
QY 178 DLDNRNTPLMSSRPHPNMYLP-IFMHGKPNRSP-NTPSHEARQPTNEFFRAPELKFQYV 235
DB 136 DLSANKLNTLTSNQSSEVSLSLDFEFTFNSDKNQPLKLPFGSVVAVANE-----SYVYSVK 191
QY 236 VVYKAAEDLMGSDIMFGYTOOSHMOIFNGKNSRFRHDYQPEL-FLTOYVYSL-PW 293
DB 192 ATLQKLKVLTYSRADHSVIGISYAIPTVSLNGKTQNDSEFNPKSNINFAKKNVYNALNPF 251
QY 294 DGVKRMIGMAVHNS--NGESAK-----LSRSWNRAYLMAGEMKNTLVMPRIMGRIK 345
DB 252 EAQOYLIVGGKFLNCKVKNADVDKNDINNHLETOFVNAKITA-----TLGKAKK 300
QY 346 ---EGSGSQP-----DNDPDILDY--GYGDV-----RFLYQL-- 373
DB 301 OFGEHKNQGPLSLKLKVLGSLNNEFKOLFVYVRPGLDLDFVSDLIQSSSGSNKKTYYQLLF 360
QY 374 ENKSN 379
DB 361 ENKTYI 366

Search completed: November 30, 2001, 14:17:45
Job time: 222 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:10:38 ; Search time 72.04 Seconds

(without alignments)
434.475 Million cell updates/sec

Title: US-09-787-083-2

Perfect score: 2360
Sequence: 1 MKVSLTLTLTLSTLSCFAILA.....YNHEATSPFGVGLMNDKMG 442

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

A_Geneseq_1101.*
1: /SID88/gcgdata/geneseq/geneseqp/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseqp/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseqp/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseqp/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseqp/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseqp/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseqp/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseqp/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseqp/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseqp/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseqp/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseqp/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseqp/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseqp/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseqp/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseqp/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseqp/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseqp/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseqp/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseqp/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseqp/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseqp/AA2001.DAT.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2360	100.0	442	21	BA5B034 amino acid
2	2354	99.7	442	21	BA5B034 amino acid
3	2352	99.7	442	21	BA5B034 amino acid
4	2330	98.7	442	21	BA5B034 amino acid
5	821	34.8	370	21	Neisseria gonorrhoe
6	814	34.5	370	21	Neisseria meningit
7	814	34.5	370	21	Neisseria meningit
8	814	34.5	374	21	Neisseria meningit
9	811	34.4	375	21	Neisseria meningit
10	246.5	10.4	355	19	H. pylori GHP0 172
11	239.5	10.1	356	19	H. pylori ORF 07ap

12	156.5	6.6	253	18	AAW20760
13	117	5.0	1686	9	AAW70991
14	110.5	4.7	800	18	AA70420
15	110.5	4.7	822	13	AA26021
16	109	4.6	824	21	AA23180
17	107	4.5	157	18	AAW20538
18	103	4.4	798	21	AAW40925
19	103	4.4	798	21	AAW90225
20	103	4.4	798	22	AAW83944
21	102.5	4.3	537	22	AAW91000
22	100.5	4.3	1726	18	AAW38756
23	98.5	4.2	467	20	AAW13378
24	98.5	4.2	467	21	AAW94860
25	98.5	4.2	467	22	AAW88401
26	98.5	4.2	467	22	AAW80246
27	96	4.1	888	22	AAW07051
28	95.5	4.0	502	22	AAW90031
29	95.5	4.0	516	22	AAW78876
30	95.5	4.0	682	17	AAW04359
31	95	4.0	1227	22	AAW81501
32	94	4.0	372	16	AAW69607
33	94	4.0	522	22	AAW95616
34	93.5	4.0	467	19	AAW75057
35	93	3.9	761	20	AAW99084
36	93	3.9	1717	22	AAW20498
37	92	3.9	600	22	AAW75092
38	92	3.9	1627	16	AAW67538
39	91.5	3.9	857	13	AAW2814
40	91.5	3.9	1024	18	AAW19604
41	91	3.9	564	21	AAW10457
42	91	3.9	564	21	AAW81952
43	90	3.8	522	22	AAW95513
44	90	3.8	790	22	AAW93045
45	90	3.8	841	20	AAW24318

ALIGNMENTS

RESULT 1
ID AAY85268 standard; Protein; 442 AA.
XX AAY85268;
AC
XX
DT 29-JUN-2000 (first entry)
XX
DE BASB034 amino acid sequence #1.
XX
KW Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KW vaccine; bacteriostatic; treatment; otitis media; pneumonia;
KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW hearing loss; antibacterial drug.
XX
OS Moraxella catarrhalis.
XX
PN WO200015802-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99WO-EP06781.
XX
PR 14-SEP-1998; 98GB-0020002.
XX
PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruelle J;
XX
DR WPI: 2000-271440/23.
XX
PT N-PSDB; AAA10700.
PT Novel BASB034 polynucleotides and polypeptides from Moraxella
catarrhalis used to prepare vaccines against bacterial infections

XX Claim 3: Fig 2: 106pp: English.

XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 CC strain Mc2931 (ATCC 43617). The invention relates to BASB034 polypeptides
 CC from M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The
 CC BASB034 polynucleotides and polypeptides may be employed as research
 CC reagents and material for the discovery of treatments and diagnostics for
 CC diseases, particularly human diseases. They are particularly used to
 CC diagnose and treat M. catarrhalis infections. They can be used for
 CC diagnosis of disease, staging of disease, or determining response of an
 CC infectious organism to drugs. The polynucleotides may be used as a source
 CC for hybridization probes, and for screening of genetic mutations,
 CC serotype, organism or strain identification, identification of mutations
 CC in BASB034 sequences, and as components of arrays which are useful for
 CC diagnostic and prognostic purposes. The polypeptides can be used to
 CC produce antibodies. The polypeptides can also be used in vaccine
 CC formulations, and to identify agonists and antagonists. The polypeptides,
 CC antibodies, agonists and antagonists (which are bacteriostatic) are used
 CC for the treatment and prevention of diseases such as otitis media in
 CC infants and children, pneumonia in elderly, sinusitis, nosocomial
 CC infections and invasive diseases, and chronic otitis media with hearing
 CC loss. The polypeptides, agonists and antagonists are also used for
 CC screening of antibacterial drugs. The BASB034 products of the invention
 CC can be used screen for new antibacterial compounds that may target
 CC resistant bacteria.

SO Sequence 442 AA:

Query Match 100.0%; Score 2360; DB 21; Length 442;
 Best Local Similarity 100.0%; Pred. No. 4.5e-219;
 Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSLSTLTLSLSCFPAITAOAOAVNPVAFVDEVSENDLGDNEPLDVOSATQSA 60
 DB 1 mkvslstltlsisclfaiaiqagavnpvafdevrsendlgdnepldvgsatqsa 60
 QY 61 STDTPANPLDEHEPELYTTALEENKTMNLINCSALNODIMRLACYDTLVHGEPVAVITKRSI 120
 DB 61 stdtampndehpeelyttalenkcmllncsalngldimrlacydtlvhgcpaviktksrsl 120
 QY 121 RLDETITWOTIKGKPOVIVQETTPFLMGNEKGMILTKKDAQOLEYAAQFTPLSFPDL 180
 DB 121 rldetltwltikgkpvivqettpflmgnekglmlltkkdaqlleyaakqfplslsfld 180
 QY 181 RNNPPLSSRRHNPVYLPFIMHGKPNRSPNTPSHAKQFTPNRPAPELKFQYSVKKA 240
 DB 181 rnnpplssrrhnpmyvlpfimbhkpnrspntpsheakqfptnreapelkfqsavkka 240
 QY 241 AEDLMGTDSDLMFGYTOQSHWQIENGKNSRPFVHDXQPEIFLTPQPVYSDLPMGDKVMI 300
 DB 241 aedlmgtddslmfgytoqshwqiengknsrpfvhdxyqpeifltqpvysdldpmgdkvmi 300
 QY 301 GMGAVHNSNGSASAKLSRSMNRAIYLMAGMEKNLYVPRIMGRIFREGSGSQPDNDPDL 360
 DB 301 gmgaavhnsngesaklsrsmnraylmagmewknlvmpriwgrlfregsgsqpddndpdlld 360
 QY 361 YVYGDVAFELVLEKNKNSIGTVRNPRSGKALQDLVYVPLGKIGSFQFOFGYGSOL 420
 DB 361 yvygdvafelyleknknsigtvrynprsgkalqldlyvyplogkisygfqlfgygsol 420
 QY 421 IDYNHEATSPFGVGLMLNDMNGL 442
 DB 421 idynheatspfgvglmndmngl 442

RESULT 2
 ID AAY85271 standard; Protein: 442 AA.
 AC AAY85271;
 XX

DT 29-JUN-2000 (first entry)
 XX BASB034 amino acid sequence #4.

XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;
 KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KW hearing loss; antibacterial drug.

XX Moraxella catarrhalis.

PN WO200015802-A1.

PD 23-MAR-2000.

XX 14-SEP-1999; 99WO-EP06781.

PR 14-SEP-1998; 98GB-0020002.

PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

PI Ruelle J;

DR WPI: 2000-271440/23.

DR N-PSDB: AAA10703.

PT Novel BASB034 polynucleotides and polypeptides from Moraxella
 PS catarrhalis used to prepare vaccines against bacterial infections -
 XX Claim 3: Page 69: 106pp: English.

CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 CC strain Mc2969. The invention relates to BASB034 polypeptides from
 CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
 CC polynucleotides and polypeptides may be employed as research reagents and
 CC material for the discovery of treatments and diagnostics for diseases,
 CC particularly human diseases. They are particularly used to diagnose and
 CC treat M. catarrhalis infections. They can be used for diagnosis of
 CC disease, staging of disease, or determining response of an infectious
 CC organism to drugs. The polynucleotides may be used as a source for
 CC hybridization probes, and for screening of genetic mutations, serotype,
 CC organism or strain identification, identification of mutations in BASB034
 CC sequences, and as components of arrays which are useful for diagnostic
 CC and prognostic purposes. The polypeptides can be used to produce
 CC antibodies. The polypeptides and antagonists (which are bacteriostatic),
 CC and to identify agonists and antagonists. The polypeptides, antibodies,
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elderly, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.

SO Sequence 442 AA:

Query Match 99.7%; Score 2354; DB 21; Length 442;
 Best Local Similarity 99.5%; Pred. No. 1.7e-218;
 Matches 440; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKVSLSTLTLSLSCFPAITAOAOAVNPVAFVDEVSENDLGDNEPLDVOSATQSA 60
 DB 1 mkvslstltlsipcfailaiaiqagavnpvafdevrsendlgdnepldvgsatqsa 60
 QY 61 STDTPANPLDEHEPELYTTALEENKTMNLINCSALNODIMRLACYDTLVHGEPVAVITKRSI 120
 DB 61 stdtampndehpeelyttalenkcmllncsalngldimrlacydtlvhgcpaviktksrsl 120
 QY 121 RLDETITWOTIKGKPOVIVQETTPFLMGNEKGMILTKKDAQOLEYAAQFTPLSFPDL 180
 DB 121 rldetltwltikgkpvivqettpflmgnekglmlltkkdaqlleyaakqfplslsfld 180

QY 181 RANPTPLMSSRPNNPMYVLPDIFMHGKPNRSPNTPSHEAKOFTPENEPAPLKTQVSVKVA 240
DB 181 rnmclwssrphnmpvlpdifmhgkpnrsnpncpsneakqfcpneitapelktqvsykva 240
QY 241 AEDLMGTSDDLWFGYTOOSHMQIFNGKNSRPFVRHDYQPEIFLTQPVYSDLPMDGKVRMI 300
DB 241 aedlwgtddslwfytqgshwqifngknsrpfvrhdypelfltqpvysdlpmdgkvrm 300
QY 301 GNGAVHNSGESAKLSRSNNRAYLMAGMWKMLTWPRIMGRIFKRGSSQPDNDPDIID 360
DB 301 gngavhnsgeesaklsrswraylmagmewknlvmpriwgrlfxegsgsqdpddndpdiid 360
QY 361 YGYGDVPRFLYQLEKNSNIGTVRNPFRSGKALQLDYVYPLGKISGYFOIFOGYGOST 420
DB 361 ygygdvprfllyqlenknsnigstvrynprsgkalqldyvyplygklsygfqifqy9gs1 420
QY 421 IDYNHEATSFVGGLMNDMWGL 442
DB 421 idynheatsfyvgllmndwmgl 442

RESULT 3
AA85270 standard; Protein: 442 AA.
AY85270:
29-JUN-2000 (first entry)
BASB034 amino acid sequence #3.
Moraxella catarrhalis infection; BASB034; diagnosis; staging;
vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
sinusitis; nosocomial infection; invasive disease; chronic otitis media;
hearing loss; antibacterial drug.
Moraxella catarrhalis.
WO200015802-A1.
23-MAR-2000.
14-SEP-1999; 99WO-EP06781.
14-SEP-1998; 98GB-0020002.
(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
Ruelle J;
WPI: 2000-271440/23.
N-PSDB: AAA10702.
Novel BASB034 polynucleotides and polypeptides from Moraxella
catarrhalis used to prepare vaccines against bacterial infections
Claim 3; Page 68; 106pp; English.

This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
strain Mc2913. The invention relates to BASB034 polypeptides from
M. catarrhalis strains Mc2913, Mc2908, Mc2913 and Mc2959. The BASB034
polynucleotides and polypeptides may be employed as research reagents and
material for the discovery of treatments and diagnostics for diseases,
particularly human diseases. They are particularly used to diagnose and
treat M. catarrhalis infections. They can be used for diagnosis of
disease, staging of disease, or determining response of an infectious
organism to drugs. The polynucleotides may be used as a source for
hybridization probes, and for screening of genetic mutations, serotype,
organism or strain identification. Identification of mutations in BASB034
sequences, and as components of arrays which are useful for diagnostic
and prognostic purposes. The polypeptides can be used to produce
antibodies. The polypeptides can also be used in vaccine formulations,

CC and to identify agonists and antagonists. The polypeptides, antibodies,
CC agonists and antagonists (which are bacteriostatic) are used for the
CC treatment and prevention of diseases such as otitis media in infants and
CC children, pneumonia in elders, sinusitis, nosocomial infections and
CC invasive diseases, and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.
XX
SQ Sequence 442 AA:

Query Match 99.7%; Score 2352; DB 21; Length 442;
Best Local Similarity 99.3%; Pred. No. 2,7e-218;
Matches 439; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSLSTLTLSTLSCPAIIAIOQAQAVPNPVAFVDEVRSENDIGDNELEPIDVQSAQSA 60
DB 1 mkvslstlslstlscfaiilaigqakavnpvatvdevrsendlgdnelpidvgsatgsa 60
QY 61 STDANPDLDEHEPELYTTALENTKMLINCSALNQDIMRLACYDTLVHGERPAVYIKRKS1 120
DB 61 stdanpldehepelytlaenktmlincsalnqdimrilaacydtlvhgepaviktkrs1 120
QY 121 RUDETIWOTIKGKPOYIYQETTPDIFLMGNEKGLRKKAQKQLEVAKQFTPLSLFDD1 180
DB 121 rldetlwtqtklkgpyvygectdpilfmgnekmltkkakaqleyakqftpls1fdd1 180
QY 181 RANPTPLMSSRPNNPMYVLPDIFMHGKPNRSPNTPSHEAKOFTPENEPAPLKTQVSVKVA 240
DB 181 rnmclwssrphnmpvlpdifmhgkpnrsnpncpsneakqfcpneitapelktqvsykva 240
QY 241 AEDLMGTSDDLWFGYTOOSHMQIFNGKNSRPFVRHDYQPEIFLTQPVYSDLPMDGKVRMI 300
DB 241 aedlwgtddslwfytqgshwqifngknsrpfvrhdypelfltqpvysdlpmdgkvrm 300
QY 301 GNGAVHNSGESAKLSRSNNRAYLMAGMWKMLTWPRIMGRIFKRGSSQPDNDPDIID 360
DB 301 gngavhnsgeesaklsrswraylmagmewknlvmpriwgrlfxegsgsqdpddndpdiid 360
QY 361 YGYGDVPRFLYQLEKNSNIGTVRNPFRSGKALQLDYVYPLGKISGYFOIFOGYGOST 420
DB 361 ygygdvprfllyqlenknsnigstvrynprsgkalqldyvyplygklsygfqifqy9gs1 420
QY 421 IDYNHEATSFVGGLMNDMWGL 442
DB 421 idynheatsfyvgllmndwmgl 442

RESULT 4
AA85269 standard; Protein: 442 AA.
AY85269:
29-JUN-2000 (first entry)
BASB034 amino acid sequence #2.
Moraxella catarrhalis.
WO200015802-A1.
23-MAR-2000.
14-SEP-1999; 99WO-EP06781.

Moraxella catarrhalis infection; BASB034; diagnosis; staging;
vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
sinusitis; nosocomial infection; invasive disease; chronic otitis media;
hearing loss; antibacterial drug.

Db	21	lqcaalcdnvrclacyrlrtlaaqlprssaaqggsgskvnlhtetvrrslldkgaavltvek	80
Qy	141	TTDFIFLMGNEKGMILTKKDAQLEYAKQTFPPSLSDLDLRNN-TPLMSSRPINPMVLPR	199
Db	81	ggdai-----pdsagetdlytptslmwdlkdndrlgllayrethmymlymp	127
Qy	200	IFMIGKPNRSPNTPESH-EAKQTFPNERPAELKFQVSVKKAEDMLGTSDDLMFGYTQO	258
Db	128	fwmnspnyapbspstlrgtvtgekgqgkqreclkqvsfkakieenlftktraadlwfygtqr	187
Qy	259	SHMOIFN-GKNSRPFKVHVDYQPELITLQPYSDLPWDGKYRMTGMAVHNSGESAKLRS	317
Db	188	sdwjllyngqxtksapfrntdykrellflqpvkadrfggrlrmjagfvlvhsnsgsqsrpser	247
Qy	318	SWNFAVILMAGEMKKNLVMPIRPIKESGSGSDPDNPDILIDYGVGDVRLFYOLENKS	377
Db	248	swmilyamagewekltcivprvwtafdq-sgdk-mnnpdiadymgydskldygrlndrq	305
Qy	378	NISGTVRYNPRSGKGLQDLDDYVPLGKISGYPQILFOGQSSLDLYNHNHETSGVGLMLN	437
Db	306	nyvsvlrynpkrtlygaaleaayltfrikbklkvgrvgrfhnsgyesllidynhknqslgijlmfn	365
Qy	438	DMWGL 442	
Db	366	dwdgl 370	

RESULT	6
AAV75157	
ID	AAV75157 standard; Protein; 370 AA

DT 21-MAR-2000 (first entry)

DE *Neisseria meningitidis* ORF 582 protein sequence SEQ ID NO:1788.

KM *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;
 KM antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
 KM antibacterial; gene therapy.

05 *Neisseria meningitidis*.

PN W09957280-A2.

PD 11-NOV-1999.

PF 30-APR-1999; 99WO-US09346.

PR 01-MAY-1998; 98US-0083758.

PR 02-SEP-1998; 98US-0098994.

PR 09-OCT-1998; 98US-0103749.

PR 09-OCT-1998; 98US-0103796.

XX

PA (GENO-) INST GENOMIC RES.

PI Fraser C, Galeotti C, G

PI Tetteln H, Venter JC;

DR WPI; 2000-062150/05.

XX

PT vaccines and diagnostics -

PS Claim 2; Page 904; 1453pp; English.

XX AA253015 to AA254536, AA254577 to AA254615, and AA774253 to AA775941
CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
CC and polypeptides. AA254537 to AA254576 and AA254616 to AA255473 repre-
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC *Neisseria* bacteria (e.g. meningitis and septicemia), to detect the
CC presence of *Neisseria* bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
XX may also be used in gene therapy protocols.
XX
SQ Sequence 370 AA;

Query Match	34.58;	Score 814;	DB 21;	Length 370;
Best Local Similarity	44.48;	Pred. No. 5.9e-70;		
Matches 162;	Conservative 68;	Mismatches 111;	Indels 24;	Gaps 8

```
QY      87   ICSLNDIMDLAEPDYLHGEFAV-----TKRSIRIDETIWTOTI-GRKQVITVE 140
        : : ||| : | | | | | | | | | | | | | | | | | | : : : :
Db      21   lqcaaltnvtrlaeydrilfaaqipsasgagsgskavlnltelvrslldkgeavivex 80

QY      141 TTDPTFLMGNEKGMLTKRKDAQLAEVAKOFPPLISPIPDENN--TFLSSSRPHHPMVL 190
        : : : : : : : : : : : : : : : : : : : : : :
Db      81 ggdal-----pdsasgetaalypisltamplakndrlrgllgvrehmpmjlmp 120
```

QY 200 IFMHGKPNRSPNTPSH-EAKQFTPNEFRAPELKFQVSVKKAAEDLMGTDSDLNFGYIQQ 256

Db 128 lwynspnyapgsptrgttvqekifgqkraetklqvsfksklaedllfkttradlwigtgtqr 187

259 SHMOIEN-GKNSRPFVRVHDYQPEIFLTQPVYSSDLPMGDKVRMIGMGAVHHSNGESAKLSK 31/

Db 188 sawqlynggrksaptrnidykpeitltqpvkadlptggrlrmkgagivhngsnqsrpestr 24/

QY 318 SWNRAYLMAGMEWKNLTVMPRIMGRIEKEGSGSQPDDNPDILBYYGGDVRELYQLENKS 3 /

Db 248 swrriyamagmewgkltvliprvwraiaq-sgak-nanpdaaymngygavklyrlnarq 303

3/8 NISGIVRNPRSGKALQDI VPELGKISGFQIFGIGQSLIDINHEAISFGVEMEN 45/

DD 306 nVysvILyHprCgYgaLeaayLrPrkKlKgvNLgLnYgesLdYmKkqngIgiImln 307

QY 438 DMMGL 442

DB 360 01ag1 3/0

RESULT 7
AAY75158

AC AAY75158;

DT 21-MAR-2000 (first entry)

DE *Neisseria meningitidis* ORF 582 protein sequence SEQ ID NO:1790.

KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;

KW antibacterial; gene therapy.

05 *Neisseria meningitidis*.

PN W09957280-A2.

PD 11-NOV-1999

PF 30-APR-1999; 99WO-US09346.

PR 01-MAY-1998: 980S-0083758.
 PR 31-JUL-1998: 980S-0094869.
 PR 02-SEP-1998: 980S-0098994.
 PR 02-SEP-1998: 980S-0099062.
 PR 09-OCT-1998: 980S-0103749.
 PR 09-OCT-1998: 980S-0103794.
 PR 09-OCT-1998: 980S-0103796.
 PR 25-FEB-1999: 990S-0121528.
 (CHIR) CHIRON CORP.
 (GENO-) INSTR GENOMIC RES.
 Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 Petersen J, Pizzzi M, Rappuoli R, Ratti G, Scalatà E, Scarselli M;
 Tettelin H, Venter JC;
 MPI: 2000-062150/05.
 N-PSDB: AAZ53920.
 Novel Neisserial polypeptides predicted to be useful antigens for
 vaccines and diagnostics -
 Claim 2: Page 905; 1453pp; English.
 AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAZ75941
 represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
 and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
 PCR primers used in the exemplification of the present invention. The
 polypeptides, the polynucleotides, antibodies and compositions of
 the invention can be used as vaccines, as diagnostic reagents, and as
 immunogenic compositions. The polypeptides can be used in the
 manufacture of medicaments for treating or preventing infection due to
Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
 presence of *Neisseria bacteria*, or to raise antibodies. They may also
 be used to screen for agonists or antagonists, which may themselves
 have use as antibacterial agents. The polynucleotides of the invention
 may also be used in gene therapy protocols.

Query Match	34.5%	Score 814	DB 21	Length 370
Best Local Similarity	44.4%	Pred. NO. 5.9e-70		
Matches 162	Conservative 68	Mismatches 111	Indels 24	Gaps 8
Qy	87	INCSALNODIRKLACYDYLVLHGEPYAVI----	KRRSRILDEETIMQRI-KGKPOIYOE	140
Db	21	lqcaalcldnverclacaydrilfaaqdpsasgqegqeska	inlvetrvrsldkgeayvlvke	80
Qy	141	TTDPILFLMGNEKGMILTKRKDAKOLEVRAAOF	FTPLSLSFOLDRRNN-TPLMSSRRHNPMYALP	199
Db	81	ggdal-----padsagetalylrslsmydtkdndklqgl	lgvtenpmnylmp	127
Qy	200	IFMHGKPNRSNPETPSH-EAKOFTPNEFRAPELK	FPQVSVVKAAEDLMGSDSLMEFYTOQ	258
Db	128	lwynsprlyagrsprcrgltvgekfsgqqrtaeklk	svsfkskaedlfkrcrdalwfsygtqr	187
Qy	259	SHMQJFN-GKNSRRPRVNDYQPELFTLPQVYS	SDLPMWCGKVRNITGMCAVHNHSGESAKLSR	317
Db	188	sdwqlynggrksapfirncdykpeellfcpvkadlp	ifggirlm1lgagfvdpqsnqgspear	247
Qy	318	SMNRAYELMAGEMWKLVTWMPRIWGRIFKESG	SSODPDNDIIDDYGGYGVRFYOLENKS	377
Db	248	swntilyamaagmewglvtrvlyprvwvafqg	-gsdk-nndpndldyngmygdvklqyrtindqg	305
Qy	378	NISGTVRYNPNRSGKALDLDYVYPLGKGISGY	FIQIQGQSGSLIDYNHEATSGVGLMNL	437
Db	306	nuyavslrynpktygataeaaylfprlkgykr	lkyvgvrfhbygseallidynhkqnglqglmfh	365
Qy	438	DMNGI 442		
Db	366	dlqgl 370		

XX	RESULT	8
XX	AAV70629	
ID	AAV70629	standard; Protein; 374 AA.
AC	AAV70629;	
XX		
DT	18-JUL-2000	(First entry)
XX		
DE	Neisseria meningitidis serogroup B strain H44/76 BASB033 protein.	
XX		
KW	BASB033; diagnosis, prophylaxis; treatment; antibacterial; vaccine;	
KW	Neisseria meningitidis infection.	
XX		
OS	Neisseria meningitidis.	
XX		
PN	W0200015801-A1.	
XX		
PD	23-MAR-2000.	
XX		
PF	09-SEP-1999; 99WO-EP06718.	
XX		
PR	14-SEP-1998; 98GB-0020003.	
XX		
PA	(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.	
XX		
PI	Ruelle J;	
XX		
DR	WPI; 2000-271439/23.	
DR	N-PSDB; AA252134.	
XX		
PT	Isolated BASB033 polypeptides and polynucleotides of Neisseria	
PT	meningitidis, useful for diagnosis, prophylaxis and treatment of N.	
XX	meningitidis infection -	
XX		
PS	Claim 4; Page 59; 93pp; English.	
XX		
CC	The present sequence is a BASB033 protein from	
CC	Neisseria meningitidis serogroup B strain H44/76. The protein	
CC	shows homology to the Klebsiella pneumoniae outer membrane	
CC	phospholipase A. The present sequence is useful for diagnosis,	
CC	prophylaxis and treatment of N. meningitidis infection. It may also be	
CC	used for the discovery and development of antibacterial compounds and	
XX	in vaccine compositions.	
XX		
XX	Sequence 374 AA;	

```

Query March 34.5%; Score 814; DB 21; Length 374;
Best Local Similarity 44.4%; Pred No. 66-70;
Matches 16; Conservative 68; Mismatches 11; Indels 24; Gaps

QY      87  INCSALNODIMRLACSYDTLVEHGETPAVI-----KTKRSIRLDETIWOTI-KGKPOVIOE 140
      : : | | : : | | | | : : : : : : : : : : : : : : : : : :
Db      25  lqcaaldntvrltlaqyrlfaaglrpsasagqegsxnvlnlteuysrsldkeavllvek 84

QY      141 TTDRIFLMGMEKMLTKKDAKOLEYAKOFTPLSLTSDFDLRNN-PLWSSRRPNMYLVR 199
      : : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      85  ggda1-----padsagetadlytrlsldyldkldnlglllgvtehmymlymp 131

QY      200 IFEMGKRNBSPTNRSH-EAKOFTNFEPRARLEKFOYSVYVUKAERDLMGSDSLMGTYOO 258
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      132 lwynspuyarpagpcttgcgtvtvcekfsgqkfiaetcklqvstfskllaedlfktrda1wfyguyqr 191

QY      259 SHMOIFN-GKNSRPFYVNDYOREPITLTOPYUSSLBPMDCIKVIMIGCAVHNSGSEAKLSR 317
      | | | | : : | | | | : | | | | | | | | | : : | : | : | : | |
Db      192 sdwqilyugqtksarptlntdykretfltpykdalrpfgrl1mlvgafvlyvsnsgsgtpear 251

QY      318 SWNRAYLMAGMEKMLTVMBRIGRIFKEGSGSQRPDNDPDLIDYUGGVRELYOLEENKS 377
      | | | | | | | | | | : : : : : : : : : : : : : : : : : :
Db      252 swmrlymagmewskltvlrgrvwvrafddq-sgk-nndprdiadumgyrvgvkllyqrlndrq 309

QY      378 NISGTVKRNPSRGALQDLDTYVRLPGKGISGVIQTGOGSGOSLIDYNNHATSFVGGLIN 437

```

QY 259 SHWC

Db 69 kky1nmndy1gtyf1

CC sequence of H. pylori (ATCC 55679) was determined from overlapping
CC contigs generated by mechanically shearing the bacterial DNA. The
CC sequences were analysed for ORF or at least 180 nucleotides, and the
CC predicted coding regions defined by computer evaluation. To identify
CC likely H. pylori antigens for vaccine development, the amino acid
CC sequences predicted from various ORF were analysed for significant
CC homology to other known or exported membrane proteins. Having identified
CC and determined the sequences of interest, particular regions can be
CC isolated from H. pylori by PCR amplification for recombinant polypeptide
CC production, e.g. in E. coli hosts.

CC
XX
SQ Sequence 253 AA:

Query Match 6.6%; Score 156.5; DB 18; Length 253;
Best Local Similarity 25.6%; Pred. No. 9.6e-07;
Matches 56; Conservative 26; Mismatches 72; Indels 65; Gaps 8;

QY 157 KKDAKOLEYAAKQFPLSLFSDLRNNTPLMSSRPNNPMYVLFPMHGKNSRPNSHE 216
DB 72 kKylmmadylgtyfLpfyhsf-----trpfqwyhpnlnp----- 105
QY 217 AKQFTNFRAPRLKQVSVKVAEDLNGTSDLMFQGYQOSHWOJFNKNSRPERVHD 276
DB 106 ---yqrnef-----kfqisfrvprfhllwtkgctlylayrqtntwfdqyndpsqpmnm 157
QY 277 YQPEITLQPVYSDLPMDCQV---RMIGCAVHHSNG-ESAKLSRSNNRAYLMAGMEKN 332
DB 158 fmpellyvypI-nfkpfqgkkgfseIwlgwqhlshngvgagcyqpfh----- 204
QY 333 LTVMPRIWGRIFKEGSGSQ--PDDNPDIIDYGYGVDYRF 369
DB 205 -----kegnpeqfpgqpvlykdyngqkdavrw 231

RESULT 13
AAW70991
ID AAW70991 standard; Protein; 1686 AA.

XX
AC AAW70991;

XX
DT 19-OCT-1998 (first entry)

XX
DE Human class II PI3 kinase-C2alpha.

XX
KW Human; class II phosphoinositide lipid kinase; PI3 kinase;

KW PIK3-C2alpha; class II; resistance; Wortmannin; LY294002.

XX
OS Homo sapiens.

XX
PN WO9832864-A2.

XX
PD 30-JUL-1998.

XX
PF 27-JAN-1998; 98MO-GB00244.

XX
PR 28-JAN-1997; 97GB-0001652.

XX
PA (LUDW-) LUDWIG INST CANCER RES.

XX
PI Domln J, Waterfield MD;

XX
DR WPI; 1998-427960/36.

XX
DR N-PSDB; AAV42920.

XX
PT New nucleic acid encoding phosphoinositol kinase 3-C2 alpha or its
PT fragments - useful for, e.g. treatment of tumour cells where
PT phenotype is associated with expression of kinase

XX
PS Claim 3; Fig 1; 52pp; English.

XX
CC The present sequence represents a human class II phosphoinositide lipid
CC (PI3) kinase designated PIK3-C2alpha. It is characterised as a class II

CC kinase due to the presence of a conserved C2 domain found in murine and
CC Drosophila class II PI3 kinases, its apparent lack of a p85 binding site
CC and a substrate affinity to inositol lipids PtdIns and PtdIns(4)P. The
CC protein has resistance to PI3 kinase inhibitors Wortmannin and LY294002.
CC Antibodies against the protein (optionally humanised), are used to
CC identify class II PI3 kinases. Antisense sequences, antibodies or
CC dominant negative mutants of the PI3-C2alpha protein, are useful in human
CC or veterinary medicine to block class II kinases. They can be used to
CC treat tumour cells where the phenotype is associated with expression of
CC PI3-C2alpha protein.

CC
XX
SQ Sequence 1686 AA:

Query Match 5.0%; Score 117; DB 19; Length 1686;
Best Local Similarity 19.2%; Pred. No. 0.13;
Matches 86; Conservative 67; Mismatches 154; Indels 140; Gaps 21;

QY 45 QDNELPIDY-----QSATQASITDTPANPLDE-----HEPELYTTALENKTMLIN----- 88
DB 522 eddetpvdlnkhlyqlepkpeamtrpveellidsynqvelalqlenqtravdqvlkav 581
QY 89 ---CSALNODIMRLACYDTLVHGETPAVIRKTSIRIDETIMQTIKGPQVITYOETDPI 145
DB 582 rklcsald-gveltal-----tesvkkklkravlnpr-----ktadv 618
QY 146 FLMGNEKGMLTRKDAKOLEYAAKQFPLSLFSDLRNNTPLMSSRPNNPMYV-----LP 199
DB 619 slfge-----dtsrst-rgslnpenpvyvsinqltaa 651
QY 200 IF-----HHGKPNRSPNPSHEAKQFTNFRAPRLKQVSVKVAARDL---NGTSDLM 252
DB 652 lydlrlrhansgrpslcaassskavkawtltetqltflf---aahgnsswnsyeky 707
QY 253 FGYYQOSHWOJFNKKN-SRPF---RVHDYQPEITLQ-----PV-YSDLPMDGKYRMI 300
DB 708 yllcsish---ngkdlfkplgskkvgtynkffyllkwdeIlfplqlsqpllesvhl 763
QY 301 GMGAVHHSNGESAKLSRSNN-----RAYLMAGMEKNLTVMPRIWGRIFK 345
DB 764 lfglngssgsspsdnkqkypaalgvkslplcdfrrflcgy-----tkllylw----- 812
QY 346 EGSQSDPDNDPDIIDYGYGVDYRFVLEENKNSISGVRNPNRSGKALQLDVYYPGKG 405
DB 813 --tsshtnsvpyvtckkgymerivlqvdfpsafdllytqpvdrlsqhmltend 870
QY 406 ISGYFOIFQGYGOSLIDYVNEHATSGV 432
DB 871 lkg-----klldllhkdsalgl 887

RESULT 14
AAP70420
ID AAP70420 standard; protein; 800 AA.

XX
AC AAP70420;

XX
DT 20-JAN-1991 (first entry)

XX
DE Sequence encoded by cellulase gene derived from Bacillus sp. No. 1139.

XX
KW Enzyme; cellotriose; cellotetrose; hydrolysis.

XX
OS Bacillus sp. No. 1139.

XX
FH Key Location/Qualifiers

FT Peptide 1..30

FT Protein 31..800

XX
PN JP6232386-A.

XX
PD 12-OCT-1987.

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:14:03 : Search time 36.79 Seconds
(without alignments)
270.358 Million cell updates/sec

Title: US-09-787-083-2
Perfect score: 2360
Sequence: 1 MKVSLSTLTLSLSCFAILA.....YNHRTSGVGLMINDMGL 442

Scoring table:
BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues.

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6C_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/6D_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101.5	4.3	1726	2	US-08-609-049A-30
2	101.5	4.3	1726	4	US-09-170-996-30
3	100.5	4.3	1658	2	US-08-609-049A-13
4	100.5	4.3	1658	4	US-09-170-996-13
5	95.5	4.0	682	3	US-08-481-435-6
6	91.5	3.9	1024	4	US-09-091-117-5
7	89.5	3.8	503	1	US-07-946-497-2
8	89.5	3.8	503	1	US-08-483-322-2
9	89.5	3.8	503	2	US-08-478-882-2
10	89.5	3.8	857	2	US-07-717-331F-2
11	87	3.7	537	1	US-08-633-879C-2
12	87	3.7	985	5	PCR-US96-03916-6
13	87	3.7	985	5	PCR-US96-03916-66
14	86.5	3.7	666	4	US-08-961-083-2
15	86	3.6	781	1	US-08-373-134D-2
16	86	3.6	781	2	US-09-114-637-2
17	86	3.6	1627	1	US-07-665-792E-9
18	84	3.6	522	6	RE34605-6
19	84	3.6	816	1	US-07-731-157A-4
20	84	3.6	816	1	US-08-229-444B-2
21	84	3.6	816	2	US-08-541-780-4
22	83.5	3.5	774	3	US-08-802-632-2
23	83.5	3.5	774	3	US-09-073-354-1
24	83.5	3.5	774	3	US-08-656-005A-1
25	83.5	3.5	774	4	US-09-073-259-1
26	83.5	3.5	774	4	US-09-363-095-1
27	83.5	3.5	774	4	US-09-418-027-1

28	83	3.5	355	1	US-07-946-497-5	Sequence 5, App1
29	83	3.5	355	2	US-08-483-322-5	Sequence 5, App1
30	83	3.5	355	2	US-08-478-882-5	Sequence 5, App1
31	82.5	3.5	535	2	US-08-564-972-1	Sequence 1, App1
32	82.5	3.5	535	4	US-09-171-969-9	Sequence 9, App1
33	82.5	3.5	657	4	US-09-306-593-2	Sequence 2, App1
34	82.5	3.5	663	4	US-08-441-139-7	Sequence 7, App1
35	82	3.5	334	2	US-08-359-850-4	Sequence 4, App1
36	82	3.5	527	2	US-08-592-126-145	Sequence 145, App
37	82	3.5	527	2	US-08-687-080-48	Sequence 48, App1
38	82	3.5	535	2	US-08-633-879C-4	Sequence 4, App1
39	82	3.5	797	4	US-09-086-912-2	Sequence 2, App1
40	82	3.5	1022	3	US-08-772-270A-2	Sequence 2, App1
41	82	3.5	1088	3	US-08-633-768A-1	Sequence 1, App1
42	81	3.4	320	2	US-08-245-511-4	Sequence 4, App1
43	81	3.4	320	2	US-08-600-993A-4	Sequence 4, App1
44	80.5	3.4	354	1	US-07-946-497-4	Sequence 4, App1
45	80.5	3.4	354	1	US-08-483-322-4	Sequence 4, App1

ALIGNMENTS

RESULT 1
US-08-609-049A-30
Sequence 30, Application US/08609049A
Patent No. 5948664
GENERAL INFORMATION:
APPLICANT: Williams, Lewis T.
APPLICANT: Moliz, Lisa
TITLE OF INVENTION: No. 5948664el PI 3-Kinase Polypeptides
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/609,049A
FILING DATE: 29-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-063700US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 1726 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-609-049A-30

Query Match 4.3% Score 101.5; DB 2; Length 1726;
Best Local Similarity 18.6% Pred No 0.47;
Matches 96; Conservative 80; Mismatches 189; Indels 161; Gaps 21;
QY 1 MKVSLSTLTLSLSCFAILAIOQAQVNPVAFVD-----EYRSNDLQGD 46
DB 465 VKVSLIEGLQLPYTFPCDVSSVTEIIMQALSHVHDLNQVDVGYTLKVCQGEVLYN 524

```

QY 47 NELPIDVOSATOSASTPANDLDEHERELTYTALENTMYLINCALMODIMRLACYTIV 106
Db 525 NHGCSHNIIONCKKWD-----EIKOLLTISAMQNLARTAEDD--- 555
QY 107 HGEPRVAIKTKRISRLDETMOITKGKROYUOETTPIF----- 146
Db 566 -EAP-----VDLKNYUOLQIKRKYKEVWTRPRVEELDSYHYVELALQOTENCHRAV 615
QY 147 -----LMGNKGLMTKKAKOLEYAA-----KOPPLISPLDNRNPLWS 188
Db 616 DOVKAANKICISALDGYETPSVTA-YKKLRAVNLPRNKSAVDYTSLSGSDTRKNSKGS 674
QY 189 SRPHNMYLVR-----IFNHGKPNR-----SPNTPSHEAKQTPNFEFRARBLKFOY 234
Db 675 INPENPVQVSMDHLTTRIDYDLRLHANSRSCSTCPRGSRNIKEAWPATE-----QL 726
QY 235 SVKKAAMDL---WGTSDMLFEGTQOSHMOIFNGKN-SRPF---RKHDOQRELFQ--- 255
Db 727 QFTYVAAHGISNNVSNEXKYLCLCSLH---NGKOLFPRIOSKKVGYLKNFEYLLKMD 782
QY 286 -----PV-YSDLPWDGKRYMIGMAVHNHNSCEAKLSRSNN-----RAYL 324
Db 783 ELIIFPIIOISQRLPESVYLHLLRFLVLNQQSSSSPDSKKOKKRGPRALOKVSLTEFDFRFL 842
QY 335 MAGMEKKNLVMPIRGIRFEGSGSQRPDDNPDILDYGYGVDFARELYOLENKSINSIGTVR 384
Db 843 TCG-----TKLILYM-----TSSHTNSIPGAIIPKRSYMERIVLYQVDFPSAPFDIY 889
QY 385 YNPRSNGALQLDVYPLGKIGISYFQIDFGVGSILIDYHNEARISFV 432
Db 890 TSPQDRNIIOODKLETFLESIDK-----KLDIDIRHSSSFL 927

```

RESULT 2
 US-09-170-996-30
 Sequence 30, Application US/09170996
 Patent No. 6291220
 GENERAL INFORMATION:
 APPLICANT: Williams, Lewis T.
 APPLICANT: Molz, Lisa
 APPLICANT: Chen, Yen-Men
 TITLE OF INVENTION: No. 6291220e1 PI 3-Kinase Polypeptides
 NUMBER OF SEQUENCES: 32
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, 8th Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-3834
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/170,996
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/609,049
 FILING DATE: 23-FEB-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Dow, Karen B.
 REGISTRATION NUMBER: 29,684
 REFERENCE/DOCKET NUMBER: 2307K-063700US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-326-2400
 TELEFAX: 415-326-2422
 INFORMATION FOR SEQ ID NO: 30:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1726 amino acids
 type: amino acid

```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-170-996-30

```

Query Match	4.38;	Score 101.5;	DB 4;	Length 1726;
Best Local Similarity	18.68;	Pred. No. 0.47;		
Matches 98;	Conservative 80;	Mismatches 189;	Indels 161;	Gaps 21

Qy	1	MKVSLIS-----TLTSLISCAIILAIQCAVNPVPVAFVD-----EVRSEINDGOD	46
Db	465	KVSLIEIGQLPVTFCDDVSSVEIITMQALMSVHDLNDQVDGVSILKVCQGEVLQN	524
Qy	47	NELPFDVOSATOSASTDTPANFLDEHEPELTYTAALEKTMALNCALNODIMRLACYTTLV	106
Db	525	NHCLGSHEHIONCKRMDI-----EIKQLTLTTSAMCONLARIYAEDD---	565
Qy	107	HGETPAVIAKTRSRIRLDETIMOTIGKRPQVYQETTPIF-----	146
Db	566	---EAP-----VDLNKLYLQIEIKPEKVEVTRHPVELLDSTHYQVELQENQHRAV	615
Qy	147	-----LMGNEKGMILYKKAQOLEYAA-----KQFTPLSLSTFDLDRNPTLMS	188
Db	616	DOVIAKAVKICSAIDQGETPVSYTEA-VKILKRAVNLPRNKSADVTSLSGSDTRNKS	674
Qy	189	SRHPNPNVLP-----IFMHGKPNR-----SPMTPSHSEAKQTPNPEFRABELKFQV	234
Db	675	LNPEHPVQVSMDLTLTRTYDLRLRHANSSRCSGCPGSSNRINKAMWATE-----QL	726
Qy	235	SVKKAAMDL---MGTSDMLFGFTQOSHWOIFNGKN-SRPE---RVHDIQPELFTQ--	285
Db	727	QFTVYAAGHIGISSNWSNVEKYYLLCSLSH---NGKDLFKFIQSKKVGTYKKNFFYLKMD	782
Qy	286	-----PV-YSDLPMDGKRYMIGMGAVHHSNESAKLRSNN-----RAYL	324
Db	783	ELIIFPIQISQPLESVYHLTLFGLVLANSSSSPDSNKKQRKGPALGKLTLEDFKRFL	842
Qy	325	MAGMEMKMLTVMPRIWGHIFKEGSGSQDPDNDPILDYGYGVDRVFLYQLEKNSNISCTVR	364
Db	843	TCG-----TKLTYLW-----TSSHTNSIPGAILPKKSYMERIVLQVDFPSPAFDIY	889
Qy	385	YNPRSGKALOLDVYVPLGKIGISYFQIFQGVGSLDINDYNEARSPGV	432
Db	890	TSPQIDRNIIOODKLETESDIKG-----KLIDIHROSSFGL	927

RESULT 3
 US-08-049A-13
 Sequence 13, Application US/08609049A
 Patent No. 5948664
 GENERAL INFORMATION:
 APPLICANT: Williams, Lewis T.
 APPLICANT: Molz, Lisa
 APPLICANT: Chen, Yen-Men
 TITLE OF INVENTION: NO. 5948664e1 PI 3-Kinase Polypeptides
 NUMBER OF SEQUENCES: 32
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, 8th Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-3834
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/609,049A
 FILING DATE: 29-FEB-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:

NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-063700US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1658 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-609-049A-13

Query Match 4.3%; Score 100.5; DB 2; Length 1658;
Best Local Similarity 18.9%; Pred. No. 0.56;
Matches 101; Conservative 79; Mismatches 194; Indels 159; Gaps 23;

QY 5 LSTLTLSILSCFALIAIQ---QAQAVNPVAFVDEVRS-----ENDLGODNE 48
DB 381 LSPVTQVRNMGENASVKSIEIEGLDLPVTFCDVSVTEIIMQALCWVHDLNQ--- 437
QY 49 LPIDVOS-----ATQASSTDFANPLDEHE--PELYTTALENKTMILNCSALNODIMRLAC 101
DB 438 --VDVGSYILKVCQGEQEVNLONNHCLSGSHEHIONCRKWDTEIKQLLTLMSAMCONLARTAE 495
QY 102 YDPLVHGEPRAVIKTRKIRLDETIMOTIKGPQVIYQETTDPIF----- 146
DB 496 DD-----EAP-----VDLNTKLYQIEKPYKEVMIHRHVEELLDSTHYQVELALOTEN 542
QY 147 -----LMGNEKMLTKKDAKOLEYAA-----KOPTPLSLSFDLDRNN 183
DB 543 QHRAVDQVIAKVRKICSLDGVETPSTEA-VKKLKRAVNLPRKKSADVTSLSGSDTRKN 601
QY 184 TPLWSSRPHNPVYLP-----IFMHGKPNR---SPNTPSHAKOFTPNREFRAPE 229
DB 602 STKGSLENPENPVOSMDHLTTATYDLRLHANSSRCSTGCPGRSNNIKEMWTATE----- 656
QY 230 LKFOVSVKVKAABDL---WGTDSDLMEGYTQOSHWOIFNGKN-SRPF---RVHDYQPEIF 282
DB 657 --QDFTVYAAHGISSNWNVSNKYKYLICLSH---NCKDLFKPIQSKKVGTYKNKFY 709
QY 283 LTO-----PV-YSDLPMDGKVRMIGMGAHVHNSGSAKLSRWN----- 320
DB 710 LKMWDELIIFPIQISQPLESVLHLTLFGVLNOSGSSPSDNKORKPEALGKVSLLTFD 769
QY 321 -RAYLAGMEMKMLTYMPRIWGRIFKESGSGQPDNDILDYGYGDVRFYOLENKSNI 379
DB 770 FKRFLLCG-----TKLLYLM-----TSSHTNSIPGALPKKSYMERIVLVQVDFPSPA 816
QY 380 SGTVRYNPRSGKALQLDVYYPPLGKISGYFOIFOGYGSGLIDYNHEATSGV 432
DB 817 FDIITYSPQIDRNIIQODKLETLESIDIKG-----KLDDIHRDSSFGL 859

RESULT 4
US-09-170-996-13
Sequence 13, Application US/09170996
Patent No. 6291220
GENERAL INFORMATION:
APPLICANT: Williams, Lewis T.
APPLICANT: Moliz, Lisa
APPLICANT: Chen, Yen-Wen
TITLE OF INVENTION: No. 6291220e1 PI 3-Kinase Polypeptides
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA

ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/170,996
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/609,049
FILING DATE: 29-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-063700US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1658 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-170-996-13

Query Match 4.3%; Score 100.5; DB 4; Length 1658;
Best Local Similarity 18.9%; Pred. No. 0.56;
Matches 101; Conservative 79; Mismatches 194; Indels 159; Gaps 23;

QY 5 LSTLTLSILSCFALIAIQ---QAQAVNPVAFVDEVRS-----ENDLGODNE 48
DB 381 LSPVTQVRNMGENASVKSIEIEGLDLPVTFCDVSVTEIIMQALCWVHDLNQ--- 437
QY 49 LPIDVOS-----ATQASSTDFANPLDEHE--PELYTTALENKTMILNCSALNODIMRLAC 101
DB 438 --VDVGSYILKVCQGEQEVNLONNHCLSGSHEHIONCRKWDTEIKQLLTLMSAMCONLARTAE 495
QY 102 YDPLVHGEPRAVIKTRKIRLDETIMOTIKGPQVIYQETTDPIF----- 146
DB 496 DD-----EAP-----VDLNTKLYQIEKPYKEVMIHRHVEELLDSTHYQVELALOTEN 542
QY 147 -----LMGNEKMLTKKDAKOLEYAA-----KOPTPLSLSFDLDRNN 183
DB 543 QHRAVDQVIAKVRKICSLDGVETPSTEA-VKKLKRAVNLPRKKSADVTSLSGSDTRKN 601
QY 184 TPLWSSRPHNPVYLP-----IFMHGKPNR---SPNTPSHAKOFTPNREFRAPE 229
DB 602 STKGSLENPENPVOSMDHLTTATYDLRLHANSSRCSTGCPGRSNNIKEMWTATE----- 656
QY 230 LKFOVSVKVKAABDL---WGTDSDLMEGYTQOSHWOIFNGKN-SRPF---RVHDYQPEIF 282
DB 657 --QDFTVYAAHGISSNWNVSNKYKYLICLSH---NCKDLFKPIQSKKVGTYKNKFY 709
QY 283 LTO-----PV-YSDLPMDGKVRMIGMGAHVHNSGSAKLSRWN----- 320
DB 710 LKMWDELIIFPIQISQPLESVLHLTLFGVLNOSGSSPSDNKORKPEALGKVSLLTFD 769
QY 321 -RAYLAGMEMKMLTYMPRIWGRIFKESGSGQPDNDILDYGYGDVRFYOLENKSNI 379
DB 770 FKRFLLCG-----TKLLYLM-----TSSHTNSIPGALPKKSYMERIVLVQVDFPSPA 816
QY 380 SGTVRYNPRSGKALQLDVYYPPLGKISGYFOIFOGYGSGLIDYNHEATSGV 432
DB 817 FDIITYSPQIDRNIIQODKLETLESIDIKG-----KLDDIHRDSSFGL 859

RESULT 5
US-08-481-435-6

Sequence 6, Application US/08481435
Patent No. 6027906
GENERAL INFORMATION:
APPLICANT: Balganes, Tanjore S
APPLICANT: Town, Christine
TITLE OF INVENTION: No. 6027906el Polypeptides
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,435
FILING DATE: 10-JUL-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IN 580/MAS/94
FILING DATE: 01-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9404072-2
FILING DATE: 24-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sterner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-151
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
FAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 682 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-481-435-6

Query Match 4.0%; Score 95.5; DB 3; Length 682;
Best Local Similarity 20.4%; Pred. No. 0.43;
Matches 92; Conservative 66; Mismatches 191; Indels 103; Gaps 24;

QY 13 LSCFAIILAIQQAQVAPNPAFVDEVRENDL-GODNELPIDVQATOS-----AST 62
DB 5 LSESKLVATTSKIDNKQNLADISSERRVNAQANDIPTDLVKAIVSIEDHREFDHKGI 64
QY 63 DTANPLDHEPELVTALNTKMLNCSALNODIMRLACDYDTLVGNETPAVIKTRIRL 122
DB 65 DTIRLGLFLRLNLSQSG-----SALTQOLIKILYFSTISDQITIS-----RK 110
QY 123 DETIWTI-----GKRPQVYIETTPDPIFLMGNEKGLTK-----KDAKQLE----- 164
DB 111 AQEAMLAQLQEGKATKQELIYIYINKVYMSNGVMOQAQNYGKDLNLSLPLALLA 170
QY 165 ---YAAKQFTPLS-LSFDDLNNTPLMSSRRPNPVTLPFIEMHGKPNNSPTPSHEAKOF 220
DB 171 GMPQAPNCPYDPSHPEDAADRNLVL--SEMKNOGYI-----SAEQYERAVNTPITDGLQS 224
QY 221 TPNERARE-----LKPOYS-VKVKAAEDLWGTSDSLWEGTYQ--OSH-WOLFNGKNSRP 271
DB 225 LKSSNRYAYADNYLKEVYINOVEETGTNLTLTGMDVYTNNDQEKHLMWDIYNDVEYA 284
QY 272 FRVHDYQPEIFLTQPVYSDLPWDGKVRMIGMGAHVHSN-----GSAKLSRSMRAYLM 325
DB 285 YPDELO-----VASTIYDVS-NGKV-IAQLGARHSSVNSGICINQAVETNDW----- 331

QY 326 AGMEKKNLT-VMPRIWGIREFGSGOPDNDLIDYQ-----YGDVRFILYQ 372
DB 332 -GSTKPTIDYPALEYGV-ESTATIHDEP--YNYPGTNPVYMDRGYFGNITLOYA 387
QY 373 LENKSNISGIVRYN-----PRSGKALQDLY 398
DB 388 LQOSRNPVAVETLTKVGLNRAKTFPLNGLIDY 419

RESULT 6

US-09-091-117-5
Sequence 5, Application US/09091117
Patent No. 6171589
GENERAL INFORMATION:
APPLICANT: The University of Melbourne
TITLE OF INVENTION: Mycoplasma Recombinant Polypeptides and
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: GREENLEE, WINNER and SULLIVAN P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: United States of America
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/091,117
FILING DATE: 12 JUNE 1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU96/00803
FILING DATE: 13-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PNT127
FILING DATE: 13-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: WINNER, Ellen P.
TELECOMMUNICATION INFORMATION:
TELEPHONE: +1 303 499 8080
FAX: +1 303 499 8089
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1024 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Mycoplasma genitalium
US-09-091-117-5

Query Match 3.9%; Score 91.5; DB 4; Length 1024;
Best Local Similarity 20.2%; Pred. No. 2.2;
Matches 86; Conservative 61; Mismatches 172; Indels 107; Gaps 22;

QY 1 MKVS-LSPTLTLSILSCFALAIQQAQVAPNPAFVDEVRENDLGDNDLPI-DVQATQ 58
DB 1 MKLSTTTTCLISGAFGTAT-----ALFTYVALKLNHQQQNTKEQON--PIKDIRGLN 54
QY 59 SASSTANPLDHEPELVTALNTKMLNCSALNODIMRLACDYDTLVGNETPAVIKTR 118
DB 55 NQVQNTIPL-HQIVVEVT--NNKAIYDKAPQKFL-----AKSALNNKL 98
QY 119 STRDETITWTIKGKRPQVYIETTPDPIFLMGNEKGLTKKAKQLEVAKO-FTPLSLSF 177
DB 99 QVEFDFKFLRT-----GVINALNADLKEWIDQTLFIPNSFF 135
QY 178 DLDKNNTPLMSSRRPNPVTLP-IFMHGKPNRSP-NTPSHEAKQTPNFRFARDELKFOVS 235

Db 136 DLSANKLNLTLNSQSEVSLDLEFIETNFSDKNQPLKLPFGSVVYVANE----SYTSVK 191
QY 236 VKKAADLMGTSDLMFGTQOOSHMOIFNGKSRPRRVADYQPEI-FLTPYVSDL-PW 293
Db 192 ATLQKRVLYTSRADSHVGSYVAIPYVSLNGKTONDFSPFKNINFAFKNYNNALNPF 251
QY 294 DGKVRIGAGVHNS--NGESAK-----LSRSMNRAVYLMAQEMKMLVMPRIKGRIFK 345
Db 252 EAOQYLVGCGKFLNOKVKNADVDKNDINNHIETQFNVAKITA-----TLGKAFK 300
QY 346 ---EGSGSQP-----DDNPILDY--GYGDV-----RFLYOL-- 373
Db 301 QFGEHNGQPLSLKVLKSLGSLNNEFKQLFNVRPGLDGFVSDLIQSSGSKKTVYQLLF 360
QY 374 ENKSN 379
Db 361 ENKTTI 366

RESULT 7
US-07-946-497-2
; Sequence 2, Application US/07946497
; Patent No. 5506119
; GENERAL INFORMATION:
; APPLICANT: HERRLICH, Peter
; APPLICANT: PONTA, Helmut
; APPLICANT: GUENTHERT, Ursula
; APPLICANT: MATZKU, Siegfried
; APPLICANT: WENZL, Achim
; TITLE OF INVENTION: VARIANT CD4 SURFACE PROTEINS, DNA
; TITLE OF INVENTION: SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
; TITLE OF INVENTION: AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/946,497
; FILING DATE: 19921109
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16915/145
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 503 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-946-497-2
Query Match 3.8%; Score 89.5; DB 1; Length 503;
Best Local Similarity 18.4%; Pred. No. 1.1;
Matches 57; Conservative 32; Mismatches 96; Indels 125; Gaps 14;

QY 100 ACYDTLVHGEPAVYIKTKRSI-----RLDETIMQTIKGRPOVYQETTPDIFLMNGEK 153
Db 213 ---DAFFIGSTLTIATTPVWVSAHTKONERTQNPVHNSPEVILQTT----- 258
QY 154 MLTKKAKOLEYAKOFTPLPSFLDRLRNTPL----WSSRPHPVYLPIFMHKGKPRS 209
Db 259 -----RMT-----DIDRNSTSAHGEWVTOEPQPPF----- 283
QY 210 PNTPSHEAKOFTPNEFPAPELKFOVSVKAAEDLW-----CTSDLMFGTQOS 259
Db 284 -NNHEYODEEETPH-----ATSTTWADPNSTTEAATQKEKF-----EN 322
QY 260 HMOIFNGKN-----SRFRVHYQPEIFLQPYVSDLPW----DGKVRMTGMG 303
Db 323 EMQ---GKNPPPSSEDSHTEGTTASAHNNHPSORMTTQSOEDVSWTDFDPISHPMGQ 379
QY 304 AVHNSGESA 313
Db 380 HOTESKCHSS 389

RESULT 8
US-08-483-322-2
; Sequence 2, Application US/08483322
; Patent No. 5760178
; GENERAL INFORMATION:
; APPLICANT: HERRLICH, Peter
; APPLICANT: PONTA, Helmut
; APPLICANT: GUENTHERT, Ursula
; APPLICANT: MATZKU, Siegfried
; APPLICANT: WENZL, Achim
; TITLE OF INVENTION: VARIANT CD4 SURFACE PROTEINS, DNA
; TITLE OF INVENTION: SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
; TITLE OF INVENTION: AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,322
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/946,497
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16915/145
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 503 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-483-322-2
Query Match 3.8%; Score 89.5; DB 1; Length 503;
Best Local Similarity 18.4%; Pred. No. 1.1;
Matches 57; Conservative 32; Mismatches 96; Indels 125; Gaps 14;

QY	40	ENDJGONDELRLDVQASATQASASTJAPRLDHEBEELVTLLEKTMJLINC.SALNOIMRL	99
Db	169	QEDIDASIMIDEDEVSS-----GSTIKRSPEEGVILMTDLPT-----SQPDRDRO-----	212
OY	100	ACYUTLVNGELPRAVIKTKRKST-----RLDETTIMQTKGKROYUYOETDPIFLMGNEKG	153
Db	213	---DAFFIGTSLAIIATTPMWSAHTKONQOETQWNPHTSNPEVLLQTTT-----	258
OY	154	MILTKKADAKOLEYAAKOFETPLISFDLORNNPTL-----WSSRPINPMVULPIFMHCKPRNS	209
Db	259	-----RMT-----DIDRNSTAGHENWTOEQDQPRF-----	283
OY	210	PNTPSHEAKOFETPNEFARPELAKFOVYSKXVLAEDLW-----CTDSDLMEFTQOOS	259
Db	284	-NNHEUYQOEELPRH-----ATSTTMDPNSTTEEAATQOREKNF-----EN	322
OY	260	HMOJFNCKN-----SRPFRVHDYQOPLEFLTOPVYSDLPW-----DGKVRMIGMS	303
Db	323	EMO-----GNRPPTPSDESHVTEGTTASAHNNHPRSQMTTQOQEDVSWYDFDPIPSHWGOG	379
OY	304	AVNHNNGESA	313
Db	380	HOTESKGHSS	389

```

1      RESULT 9
2      US-08-478-882-2
3      : Sequence 2, Application US/08478882
4      : Patent No. 5885575
5      :
6      : GENERAL INFORMATION:
7      :
8      : APPLICANT: HERRLICH, Peter
9      : APPLICANT: PONTA, Helmut
10     : APPLICANT: GUENTHERT, Ursula
11     : APPLICANT: MATZKU, Siegfried
12     : APPLICANT: WENZL, Achim
13     :
14     : TITLE OF INVENTION: VARIANT CD44 SURFACE PROTEINS, DNA
15     : TITLE OF INVENTION: SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS
16     : TITLE OF INVENTION: AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
17     : NUMBER OF SEQUENCES: 8
18     :
19     : CORRESPONDENCE ADDRESS:
20     : ADDRESSEE: Foley & Lardner
21     : STREET: 3000 K Street, N.W., Suite 500
22     : CITY: Washington, D.C.
23     :
24     : COUNTRY: USA
25     :
26     : ZIP: 20007-5109
27     :
28     : COMPUTER READABLE FORM:
29     :
30     : MEDIUM TYPE: Floppy disk
31     : COMPUTER: IBM PC compatible
32     : OPERATING SYSTEM: PC-DOS/MS-DOS
33     : SOFTWARE: Patent Release #1.0, Version #1.25
34     :
35     : CURRENT APPLICATION DATA:
36     : APPLICATION NUMBER: US/08/478,882
37     :
38     : FILING DATE:
39     :
40     : CLASSIFICATION: 435
41     :
42     : PRIOR APPLICATION DATA:
43     : APPLICATION NUMBER: US/07/946,497
44     :
45     : FILING DATE: 19921109
46     :
47     : ATTORNEY/AGENT INFORMATION:
48     :
49     : NAME: BENT, Stephen A.
50     : REGISTRATION NUMBER: 29,768
51     : REFERENCE/DOCKET NUMBER: 16915/145
52     :
53     : TELECOMMUNICATION INFORMATION:
54     : TELEPHONE: (202)672-5300
55     : TELEFAX: (202)672-5399
56     :
57     : TELEX: 904136
58     :
59     : INFORMATION FOR SEQ ID NO: 2:
60     :
61     : SEQUENCE CHARACTERISTICS:
62     :
63     : LENGTH: 503 amino acids
64     : TYPE: amino acid
65     : TOPOLOGY: linear
66     :
67     : MOLECULE TYPE: protein
68     :
69     :
70     :
71     :
72     :
73     :
74     :
75     :
76     :
77     :
78     :
79     :
80     :
81     :
82     :
83     :
84     :
85     :
86     :
87     :
88     :
89     :
90     :
91     :
92     :
93     :
94     :
95     :
96     :
97     :
98     :
99     :
100    :
101    :
102    :
103    :
104    :
105    :
106    :
107    :
108    :
109    :
110    :
111    :
112    :
113    :
114    :
115    :
116    :
117    :
118    :
119    :
120    :
121    :
122    :
123    :
124    :
125    :
126    :
127    :
128    :
129    :
130    :
131    :
132    :
133    :
134    :
135    :
136    :
137    :
138    :
139    :
140    :
141    :
142    :
143    :
144    :
145    :
146    :
147    :
148    :
149    :
150    :
151    :
152    :
153    :
154    :
155    :
156    :
157    :
158    :
159    :
160    :
161    :
162    :
163    :
164    :
165    :
166    :
167    :
168    :
169    :
170    :
171    :
172    :
173    :
174    :
175    :
176    :
177    :
178    :
179    :
180    :
181    :
182    :
183    :
184    :
185    :
186    :
187    :
188    :
189    :
190    :
191    :
192    :
193    :
194    :
195    :
196    :
197    :
198    :
199    :
200    :
201    :
202    :
203    :
204    :
205    :
206    :
207    :
208    :
209    :
210    :
211    :
212    :
213    :
214    :
215    :
216    :
217    :
218    :
219    :
220    :
221    :
222    :
223    :
224    :
225    :
226    :
227    :
228    :
229    :
230    :
231    :
232    :
233    :
234    :
235    :
236    :
237    :
238    :
239    :
240    :
241    :
242    :
243    :
244    :
245    :
246    :
247    :
248    :
249    :
250    :
251    :
252    :
253    :
254    :
255    :
256    :
257    :
258    :
259    :
260    :
261    :
262    :
263    :
264    :
265    :
266    :
267    :
268    :
269    :
270    :
271    :
272    :
273    :
274    :
275    :
276    :
277    :
278    :
279    :
280    :
281    :
282    :
283    :
284    :
285    :
286    :
287    :
288    :
289    :
290    :
291    :
292    :
293    :
294    :
295    :
296    :
297    :
298    :
299    :
300    :
301    :
302    :
303    :
304    :
305    :
306    :
307    :
308    :
309    :
310    :
311    :
312    :
313    :
314    :
315    :
316    :
317    :
318    :
319    :
320    :
321    :
322    :
323    :
324    :
325    :
326    :
327    :
328    :
329    :
330    :
331    :
332    :
333    :
334    :
335    :
336    :
337    :
338    :
339    :
340    :
341    :
342    :
343    :
344    :
345    :
346    :
347    :
348    :
349    :
350    :
351    :
352    :
353    :
354    :
355    :
356    :
357    :
358    :
359    :
360    :
361    :
362    :
363    :
364    :
365    :
366    :
367    :
368    :
369    :
370    :
371    :
372    :
373    :
374    :
375    :
376    :
377    :
378    :
379    :
380    :
381    :
382    :
383    :
384    :
385    :
386    :
387    :
388    :
389    :
390    :
391    :
392    :
393    :
394    :
395    :
396    :
397    :
398    :
399    :
400    :
401    :
402    :
403    :
404    :
405    :
406    :
407    :
408    :
409    :
410    :
411    :
412    :
413    :
414    :
415    :
416    :
417    :
418    :
419    :
420    :
421    :
422    :
423    :
424    :
425    :
426    :
427    :
428    :
429    :
430    :
431    :
432    :
433    :
434    :
435    :
436    :
437    :
438    :
439    :
440    :
441    :
442    :
443    :
444    :
445    :
446    :
447    :
448    :
449    :
450    :
451    :
452    :
453    :
454    :
455    :
456    :
457    :
458    :
459    :
460    :
461    :
462    :
463    :
464    :
465    :
466    :
467    :
468    :
469    :
470    :
471    :
472    :
473    :
474    :
475    :
476    :
477    :
478    :
479    :
480    :
481    :
482    :
483    :
484    :
485    :
486    :
487    :
488    :
489    :
490    :
491    :
492    :
493    :
494    :
495    :
496    :
497    :
498    :
499    :
500    :
501    :
502    :
503    :
504    :
505    :
506    :
507    :
508    :
509    :
510    :
511    :
512    :
513    :
514    :
515    :
516    :
517    :
518    :
519    :
520    :
521    :
522    :
523    :
524    :
525    :
526    :
527    :
528    :
529    :
530    :
531    :
532    :
533    :
534    :
535    :
536    :
537    :
538    :
539    :
540    :
541    :
542    :
543    :
544    :
545    :
546    :
547    :
548    :
549    :
550    :
551    :
552    :
553    :
554    :
555    :
556    :
557    :
558    :
559    :
560    :
561    :
562    :
563    :
564    :
565    :
566    :
567    :
568    :
569    :
570    :
571    :
572    :
573    :
574    :
575    :
576    :
577    :
578    :
579    :
580    :
581    :
582    :
583    :
584    :
585    :
586    :
587    :
588    :
589    :
590    :
591    :
592    :
593    :
594    :
595    :
596    :
597    :
598    :
599    :
600    :
601    :
602    :
603    :
604    :
605    :
606    :
607    :
608    :
609    :
610    :
611    :
612    :
613    :
614    :
615    :
616    :
617    :
618    :
619    :
620    :
621    :
622    :
623    :
624    :
```

Query	Match	Similarity	3.8%	Score	69.5	DB 2	Length	503
Match	Local	Similarity	18.4%	Pred.	No.1.1			
Batches	57	Conservative	32	Mismatches	96	Indels	125	Gaps
QY	40	ENDLQDNELPIDVOSASTDPANPLDEHPELYTALLENKTMILNCALNODIMRL	99					
Db	169	QEDIDASNIIDEDVSS-----GSTLEKSTPGYILHTDLPt-----SQPTGRDR-----	212					
QY	100	ACYDPLVNGEPAVLYKTKRST-----RLDELTINOTIGKPOVYIYQETTDPLFLMGNGK	153					
Db	213	---DAFFIGSTATATATTPWVSANHTKONDERQWNPINHSPEVLIQTTT-----	258					
QY	154	MUTKDAKQLEVAAKOFTPNELSLSPFLDRNNPTL-----WSSRHPNPMYVLPFIHGKPNRS	209					
Db	259	-----RMT-----DIDRNSTSAHGENTQOEQPP-----	283					
QY	210	PNTPSHEAKOFTPNENFEARBLKFOVSVYKKAEDLM-----GTDSDLMFGYTOOS	259					
Db	284	---NNHRYQODEEERPH-----ASTJMWADPNSTTEEAATQKEKMF-----EN	322					
QY	260	HMQITNGKN-----SRPRVADYOEITLTPQVYDLM-----DQKVMKIGMG	303					
Db	323	EMO---GKNPRTPSDESHVTEGTSASAHNNHNSQKRWITQSOEDVSMTDFEFDRIISHMGOG	379					
QY	304	AVVHNSNGESA	313					
Db	380	HOTESKGHSS	389					

```

1      RESULT 10
2      US-07-717-331F-2
3      Sequence 2, Application US/07717331F
4      Patent No. 5484905
5      GENERAL INFORMATION:
6      APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua
7      APPLICANT: Stein
8      TITLE OF INVENTION: A Receptor Protein Kinase Gene
9      TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus
10     NUMBER OF SEQUENCES: 9
11     CORRESPONDENCE ADDRESS:
12     ADDRESSEE: Yahwak & Associates
13     STREET: 25 Skytop Drive
14     CITY: Trumbull
15     STATE: Connecticut
16     COUNTRY: USA
17     ZIP: 06611
18     COMPUTER READABLE FORM:
19     MEDIUM TYPE: Floppy Disk
20     COMPUTER: Macintosh
21     OPERATING SYSTEM: MS-DOS
22     SOFTWARE: Microsoft Word
23     CURRENT APPLICATION DATA:
24     APPLICATION NUMBER: US/07/717,331F
25     FILING DATE: June 19th 1991
26     CLASSIFICATION: 800
27     ATTORNEY/AGENT INFORMATION:
28     NAME: George M. Yahwak
29     REGISTRATION NUMBER: 26, 824
30     TELECOMMUNICATION INFORMATION:
31     TELEPHONE: (203)268-1951
32     TELEFAX: (203)268-1951
33     INFORMATION FOR SEQ ID NO: 2:
34     SEQUENCE CHARACTERISTICS:
35     LENGTH: 857 amino acids
36     TYPE: amino acid
37     TOPOLOGY: linear
38     MOLECULE TYPE: peptide
39     US-07-717-331F-2

```

Query Match 3.88; Score 89.5; DB 1; Length 857;

Best Local Similarity 21.1%; Pred. No. 2.7;
Matches 75; Conservative 51; Mismatches 110; Indels 119; Gaps 23;

QY 105 LVHGETPAVITKRSIRLDETTIWTIKGKQVYIQTETDPIFLMG---NEK---GMLTK 157
 Db 24 LIRPALSIYINT-----LSSTESTLTISSNKTLV---SPGSIFFGVGFRTNSRWLGMWYK 75
 QY 158 K-DAKOLEYAKQFTPLISFD-----LDRNNTPLMSSRPINPVYLPDIFMHGK 205
 Db 76 KVSDFRYVWVANDPNLSNIGTLKISGNLVLDDSHSNKRVWMTN-----LTRGN 125
 QY 206 PNRSP-----NTPSHAKQFT-----PNEFRAPELKQVSVKYAAEDLW 245
 Db 126 -ESSPVAAELLANGNFPMRSSNNNDASEYIMOSFDPDTDLLEPMKIGYMLKT----- 177
 QY 246 GIDSIDMFEGYTOOSHMOIFGNKSRPFRVHDYQ-----PEIFLQO---PVYSDLPMDGK 296
 Db 178 GLNRFL-----TSWRSSDPPSSGNF---SYKLETSQSLPEFYLSRENFMHRSGBPWN- 226
 QY 297 VRIIGAGAVH-----HSGESAKLSRSNNRAV-----LMAGHEKMLTYMP--RI 339
 Db 227 IIRSGIPEDDKLSYMYNFIENNEEVAAYTERMTNNSFYSLTLISEGYFORLTWYPSIRI 286
 QY 340 WGRIFEGSGSGOPD-----NPDLIDYGYGVDVRELYQLEKSNISCTVR-YNPRS 389
 Db 287 WNFNRPSPDRCQDTYIMGCP-----YAYCDV-----NTSPVCNCTQGFNPRN 329

RESULT 11
 US-08-633-879C-2
 ; Sequence 2, Application US/08633879C
 ; Patent No. 5928922

GENERAL INFORMATION:
 APPLICANT: Kivirikko, Kari I.
 APPLICANT: Pihlajaniemi, Taina
 APPLICANT: Helekoski, Tarja I.
 APPLICANT: Anonen, Pia P.
 APPLICANT: Nissi, Ritva K.
 APPLICANT: No. 5928922elainen, Minna K.
 TITLE OF INVENTION: '2 SUBUNIT OF PROLYL-4-HYDROXYLASE
 TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING SUCH SUBUNIT AND
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds, LLP
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: NY
 COUNTRY: USA
 ZIP: 10036-2811

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows
 SOFTWARE: FastSEO for Windows Version 2.0b
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/633,879C
 FILING DATE: 10-APR-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:
 NAME: Adams, Samuel B
 REGISTRATION NUMBER: 30,605
 REFERENCE/DOCKET NUMBER: 8389-0041-999
 TELEPHONE: 650-493-4935
 TELEFAX: 650-493-5556
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 537 amino acids

TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FRAGMENT TYPE: internal
 US-08-633-879C-2

Query Match 3.7%; Score 87; DB 2; Length 537;
 Best Local Similarity 21.0%; Pred. No. 2.3;
 Matches 98; Conservative 53; Mismatches 146; Indels 170; Gaps 25;

QY 1 MKVSLSTLTLSISCFAILLAQQAQAVNPVAFVDESRNDGODNELPIDVQSA---- 56
 Db 1 MKLQVLVLVL-LMSWFEGLVSMVQAEFTTSIGHMTDLIYAEKDLVQSIKEYTLVEAKLAK 59
 QY 57 -----TOSASTD-----TANPLDEH-----PELYTVALENKT--MLNCS 90
 Db 60 IKSMASKMEALTSRSADPEGYLAHPVNAVYKLVRLNTDMPALGDVLDQASAGFVANLS 119
 QY 91 ALNQ-----DIMRLA-CY---DTLVHGETPAVITKRSIRLDETI----- 126
 Db 120 VORQFFPTDESGARALIRLADPTVKLDPTLSRGELPCT-KYQAMLSDDCGGLGRSA 178
 QY 127 -----WOTIKGKQVYIQTETDPIFLMGNEKMLTK-DAKOLEYAKQFTPLISFDL 179
 Db 179 YNEGDIYHTVLMMEQVLKQ-----LDAGEEATVTSVLVDYLSYAVFQGLDHRVEL 231
 QY 180 DRNNTPLMSSRPINPVYLPDIFMHGKPNRSPNPSHAK-----OFTNEFRAPELKF 232
 Db 232 TRRLSL-----DPSHERAGNLRKYFERLLEEGKSLSN 266
 QY 233 QVSVKYKAAEDLMGTDSDLMFGYTOOSHMOIFGNKSRPFRVHDYQEIFLQPVYSDLP 292
 Db 267 QTDAGLATQENLY-----RRP---FDYLP-----RDVYESL- 295
 QY 293 WDKVRIIGAGAVHHSNGESAKLS-RSMNRAIYM--AGMEKMLTYMPRIAGRIFKSGC 349
 Db 296 -----CRGEGKILPRROKRLFCRYHGNRVPOLLAP-----FKR-- 331
 QY 350 SOPDDNPDLIDYGYGVDVRELYQLEKSNIS-----GTVARNRSG 390
 Db 332 EDEWDSPIYVRY---DYMSDEIERIKETAKPKLAIATVR-DPKTG 374

RESULT 12
 PCT-US96-03916-6
 ; Sequence 6, Application PC/TUS9603916
 ; GENERAL INFORMATION:

APPLICANT: Wild, Martha A.
 APPLICANT: Cochran, Mark D.
 TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
 TITLE OF INVENTION: AND USES THEREOF
 NUMBER OF SEQUENCES: 72
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Cooper & Dunham LLP
 STREET: 1185 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US96/03916
 FILING DATE: 23-MAR-1995
 CLASSIFICATION:

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/126,597
 FILING DATE: 24-SEP-1993


```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961,083
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Brookes, A. Anders
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PB340P2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 666 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
: US-08-961-083-2

```

Query Match	3.78;	Score 86.5;	DB 4;	Length 666;
Best Local Similarity	20.78;	Pred. No. 3.6;		
Matches 88;	Conservative 57;	Mismatches 178;	Indels 103;	Gaps 24

[illegible]

RESULT 15
 US-08-373-134D-2
 : Sequence 2, Application US/08373134D
 : Patent No. 5780296
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Kmiec, Eric
 : APPLICANT: Holloman, William
 : TITLE OF INVENTION: COMPOSITIONS AND METHODS TO PROMOTE
 : TITLE OF INVENTION: HOMOLOGOUS RECOMBINATION IN EUKARYOTIC CELLS AND ORGANISMS
 : NUMBER OF SEQUENCES: 15
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Pennie & Edmonds
 :

STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/373,134D
FILING DATE: January 17, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: FriebeI, Thomas E.
REGISTRATION NUMBER: 29,258
REFERENCE/DOCKET NUMBER: 7991-007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 781 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-373-134D-2

Query Match	3.6%	Score 86;	DB 1;	length 781;
Best Local Similarity	20.7%	Pred. No. 5.3;		
Matches 50;	Conservative 37;	Mismatches 122;	Indels 32;	Gaps 9;

[illegible]

Search completed: November 30, 2001, 14:17:42
Job time: 219 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:18:32 ; Search time 44.45 Seconds
(without alignments)
757.461 Million cell updates/sec

Title: US-09-787-083-4

Perfect score: 2360

Sequence: 1 MKVSLSTLTLSTLPCFAILA.....YNHEATSFVGVLNDMMGL 442

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	815	34.5	382	2	EB1195
2	815	34.5	409	2	HB1831
3	373	15.8	286	2	B36971
4	366	15.5	329	2	D81279
5	358	15.2	289	2	A36971
6	357	15.1	289	1	PSECA1
7	357	15.1	289	2	EB6069
8	321.5	13.6	289	2	C36971
9	246.5	10.4	355	2	C64582
10	240.5	10.2	355	2	H71930
11	109.5	4.6	824	3	JC7532
12	109	4.6	800	2	A29003
13	109	4.6	822	2	JT0611
14	106.5	4.5	602	1	TVTRR
15	101.5	4.3	783	2	JC5467
16	98.5	4.2	1658	2	T42642
17	97.5	4.1	797	2	T46737
18	95.5	4.0	389	2	S76490
19	94.5	4.0	397	2	A35136
20	94.5	4.0	601	2	T26062
21	94.5	4.0	655	2	T26061
22	94.5	4.0	1641	2	D82704
23	94	4.0	696	2	S55694
24	94	4.0	765	2	T35719
25	93.5	4.0	564	2	T40777
26	93.5	4.0	901	2	T20122
27	93	3.9	633	2	C81956
28	93	3.9	791	2	H96839
29	93	3.9	4273	2	C69679

30	93	3.9	4307	2	T20721
31	92.5	3.9	857	1	A41369
32	92.5	3.9	1807	2	T30940
33	92	3.9	324	2	B69521
34	92	3.9	888	2	S50801
35	92	3.9	2971	2	T08026
36	92	3.9	5005	2	F82884
37	91.5	3.9	796	2	JC7355
38	91	3.9	367	2	T24058
39	91	3.9	1078	2	T18352
40	91	3.9	1474	1	MAHU
41	90.5	3.8	386	2	A81328
42	90.5	3.8	467	1	A49377
43	90	3.8	841	2	JC5894
44	90	3.8	1788	2	T29043
45	89.5	3.8	477	2	S23257

ALIGNMENTS

RESULT 1
EB1195
phospholipase A1, probable NMB0464 [imported] - Neisseria meningitidis (strain MC58 s
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: EB1195
R:Retelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
ri, H.; Qin, H.; Yamathavan, J.; Gill, J.; Scariato, V.; Moxon, E.R.; Rappaport, R.;
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappaport, R.;
A>Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755
A:Accession: EB1195
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-382 <TEP>
A:Cross-references: GB:AE002403; GB:AE002098; NID:q7225688; PIDN:AAFA0901.1; PID:q7222
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0464

Query Match	34.5%	Score 815;	DB 2;	Length 382;
Best Local Similarity	44.7%	Pred. No. 1.4e-59;		
Matches 163;	Conservative 67;	Mismatches 111;	Indels 24;	Gaps 8;
QY	87	INCSALNODIMRLACVDTLVHGETPAVI-----KTKRSIRLDETIWOT-KGKPOVYQE	140	
DB	33	LOCALITDVTYLACVDRFAAQLPSASQEQESKAVLNTLEYRSSLIDKEAVIYVEK	92	
QY	141	TTDPFLMGNEKGMCLKDAQLEYAARQFTPLSLFSDLRRNN-TPLWSSRPHNPVYLP	199	
DB	93	GGDAL-----PADSAGETADYITPLSLWYDLDRKDLGLCVREHNPVYLP	139	
QY	200	FMHCKPNSRPTPSHEAR-QFTNDFRAPPELKFOYSKVKAAEDLMGDSULMTCYTOQ	258	
DB	140	IMYNNSPVYAPSPGRTGVQOKRAETKLYVSEFSKJAEDLFTKRALMWCYTOR	199	
QY	259	SHMOJFN-GKNSRPRVHDOPEIFLTQPVYSDLPMDGKVRIGAGVAHNSGESAKLSR	317	
DB	200	SMOJYNOCKRKAAPRNDYKPEITLQDPVKAADLPFGGLRMAGFVHQSGQSRPESR	259	
QY	318	SNRRAVLAAGMEKMLTVAPRIWGRIFKEGSGSQDPDNDIIDDYGYGQVREPLYOLENKS	377	
DB	260	SNRRIYAAGMEKMLTVIPRVWVRAFDQ-SSDK-NDNPDIADYGYGQVRLNDROQ	317	
QY	378	NISGVIRINPRSGKALQLDIYVPLKGISGFQIPQYGGSLIDYVNEATSFVGLMIN	437	
DB	318	NYSVALRYNPKYGAIEAAVTFPIKGLKGVVRFHGGESLIDYVNHKONGIGIGLMEFN	377	
QY	438	DMNGL 442		

Db 378 DLDGI 382

RESULT 2

H81831

probable phospholipase MMA2021 [imported] - Neisseria meningitidis (strain 22491, serogroup C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: H81831
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, J.; Holtz, S.; Jørgensen, K.; Leather, S.; Mouton, S.; Mungall, K.; Quail, M.A.; Rajandream, Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491.
A:Reference number: A81775; MUID:20222556
A:Accession: H81831
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-409 <PAR>
A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85240.1; PID:g738065
A:Experimental source: serogroup A, strain 22491
C:Genetics:
A:Gene: MMA2021

Query Match

Best Local Similarity 34.5%; Score 815; DB 2; Length 409;
Best Local Similarity 44.7%; Pred. No. 1.6e-59;

Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;

QY 87 INCSALNODIMRLACDYLHGETPAVI-----KTKRSILDETIMQTT-KGKPYVYQOE 140
DB 60 LQCAALTDVTRLCYDRIFPAQLPSAGQEGESKAVLMLTETVSSLDKGEAVIVEK 119
QY 141 TDPFIFMGNEKMKTKKDAKOLEYAAKOPTPLSLSPDLDRNN-TPLMSRPNNPVLP 199
DB 120 GGAAL-----PADSAGETADITTPSLMTDLKNDLRGLGRENHMYLMP 166
QY 200 IFNHGKPNRSPNTPSHEAR-QFTPNFEFAPELKFOYSVKKAAEDLWGTSDLMFGYTOQ 258
DB 167 LMYNSPNVAPGSPRTGTVOEKFQGOAKRAETKLOVSFKIADLEFKTRADLMFGYTOR 226
QY 259 SHMOITN-GKNSRPFVNDYQPEIFLTQPYSDLPMDGKRAMIGMCAVHNSGESAKLR 317
DB 227 SDMOITNQGKSAFRTDTDKPFIPLQPKADLPFGGRILMGAGFVHNSGOSRPRESR 286
QY 318 SMNRAYLMAEMKNTLVMPRIKGRITFEKSGSOPDDNPDLIDYGGGVDFPFLYOLENKS 377
DB 287 SMNRITAMAGMEMKLTIVIPRVVRAFDQ-SGDK-NDNPIDIMYGTGDKLQTRRLMDRO 344
QY 378 NISGTVRYNPRSGKALQLDYVYPLGKGISYFOIFQYQOSLIDYNHEATSEFGVGLMLN 437
DB 345 NYSVLYRYNPKTGYGAIEAAYTFPIKGLKGVVRGFGYGESLIDYHKNKGIGIGLMFN 404
QY 438 DMWGL 442
DB 405 DLDGI 409

RESULT 3

B36971

outer membrane phospholipase A (EC 3.1.1.-) precursor - Klebsiella pneumoniae
C:Species: Klebsiella pneumoniae
C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 18-Jun-1999
C:Accession: B36971; S40129
R:Brook, R.G.P.M.; Brinkman, E.; van Boxel, R.; Bekkers, A.C.A.P.A.; Verneij, H.M.; Tomm
J. Bacteriol. 176, 861-870, 1994
A:Title: Molecular characterization of enterobacterial plda genes encoding outer membran
A:Reference number: A36971; MUID:94131966
A:Accession: B36971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-286 <PRO>
A:Cross-references: EMBL:X76901; NID:g436880; PIDN:CAA54223.1; PID:g436881

A:Note: authors translated the codon AAG for residue 112 as Arg
C:Genetics:
A:Gene: plda
C:Superfamily: bacterial phospholipase A1
C:Keywords: carboxylic ester hydrolase

Query Match

Best Local Similarity 15.8%; Score 373; DB 2; Length 286;
Best Local Similarity 42.1%; Pred. No. 3.1e-23;

Matches 90; Conservative 27; Mismatches 87; Indels 10; Gaps 5;

QY 226 RABELKFOYSVKKAAEDLWGTSDLMFGYTOOSHWQIFGKNSRPFVNDYQPEIFLTQ 285
DB 80 RKDEKVFQLSLAPPLMGRILGDSNLLGASYSQKSMWQLSNSKESAPFRETNYEPQLFGR 139
QY 286 PV-YSDLPMDGKRAMIGMCAVHNSGESAKLSMNRAYLMAEMKNTLVMPRIKGRIF 344
DB 140 ATDYQFAGW--TLRDIEMGNHDSNGRSDPTSSMNLARLMAONGNMLVEKPPW---Y 194
QY 345 KEGSGSOPDDNPDLIDYGGGVDFPFLYOLENKSINISGTVRYNPRSGKALQLDYVYPLGK 404
DB 195 VVGS-----TDDNPDTIKMYGRRLKVGYL-GEALISAGQYNNNTGCGAELGVSPITRK 250
QY 405 GISGYFOIFQYQOSLIDYNHEATSEFGVGLMLND 438
DB 251 HVRAYTOIYSYGESLIDYNFQTRGVGLMLND 284

RESULT 4

D81279

phospholipase A1 (EC 3.1.1.32) Cj1351 [imported] - Campylobacter jejuni (strain NCTC
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 18-Aug-2000
C:Accession: D81279
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chli
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; Van Vliet, A.; Whitehead, S.; Ba
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals
A:Reference number: A81250; MUID:20150912
A:Accession: D81279
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-329 <PAR>
A:Cross-references: GB:AL139078; GB:AL111168; NID:g6968723; PIDN:CAB73778.1; PID:g6966
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: plda; Cj1351
C:Keywords: carboxylic ester hydrolase

Query Match

Best Local Similarity 15.5%; Score 366; DB 2; Length 329;
Best Local Similarity 33.2%; Pred. No. 1.4e-22;

Matches 105; Conservative 48; Mismatches 125; Indels 38; Gaps 13;

QY 130 IKGRPOVYQETDPIFLMGNEKMLTKKDAKOLEYAAKOPTPLSLSPDLDRNNT--PLW 187
DB 43 LKNSVLIISOQNNSQATQONSITIKKEKQ-----DFSRLALANYGENSEFNPILG 96
QY 188 SSRPHNMYVLPFIEMHKKPRSPRTPSHEARQFTPNFEFAPELKFOYSVKKAAEDLMGT 247
DB 97 IS-SYKNNYFLP-FVYSFNSLGVNNKSEA-----KFOISVKKRLFEENLGL 141
QY 248 DSDLMFGYTOOSHWQIFGKNSRPFVNDYQPEIFLTQPY-SDLPMDGKRAMIGMCAVH 306
DB 142 DEKYIYITQTSWQIT--EHSFPFRTNYQPEFIDPLKLDYEFNNLR---VGILH 196
QY 307 HSNGESAK--LSMNRAYLMAEMKNTLVMPRIKGRITFEKSGSOPDDNPDLIDYGGY 364
DB 197 ESNKKGDENLSSMNRITVSTALYKFLFVRLMYRI--DENKKDDNDPALHYMGN 253
QY 365 GDVRFYOLENKSINISGTVRYNPR--SGKALQLDYVYPL-GKGISYFOIFQYQOSGLI 421
DB 254 FDNVLAAY-LQDDYFNTLMRLNKLKFNHNNKGAIOVDLGYDIFNNGIYIYIQLYFNGYGESL 312


```

Oy  193 NPMYVLPITPMGKPKRPSRNPNTPSHEAROTPTPMEFAPRELK-----FQVSYKKAEDLMGT 247
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  256 HTMYTVHFHYTGSNAASTSESPSE-----TPNSEKGNVMSNTRALENGVAVFATE--WGT 308
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy  248 -----DSDLMWGYTQO-----SHNQJNGKMS-----RPFVY----- 274
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  309 SQASGDGPPFYDEADVWLEFPENNISWANSLTN-KNEVSGAFTPELGKSNATNLDPG 367
Oy  275 --HDYQ-EIFTP-----OPV-----YSDLPW---DGKVRMIGVAVHSNGE 311
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  368 PDHWAPPELISGSGYVVARIKGVNYEPIDTKTKTKVLMYDNDGTKQ--GRGVNSDSPNK 425
Oy  312 SAKLSRSNNRAYLMAGME-----WKNLTWVPRIGR 342
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  426 ELIAYDNEENTLKVSGLDVSDNDSGDNFWANARLSANCMGK 466

RESULT 12
A29003
cellulase (EC 3.2.1.4), alkaline - Bacillus sp.
N:Alternate names: endo-1,4-beta-glucanase
C:Species: Bacillus sp.
C:Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 21-Jul-2000
C:Accession: A29003
R:Fukumori, F.; Kudo, T.; Narahashi, Y.; Horikoshi, K.
J. Gen. Microbiol. 132, 2329-2335, 1986

```

A; Accessio

```

A:Residues: 17800 <FUK>
A:Cross-references: GB:D00066; GB:N00066; NID:q216223; PIDN:BA00045.1; PID:q216224
A:Experimental source: Strain 1139
A:Function:
A:Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as
A:Pathway: cellulose degradation
A:Superfamily: Bacillus sp. alkaline cellulase; Thermotoga xylanase A amino-terminal
C:Keywords: glycosidase; hydrolysis; polysaccharide degradation
F:585-726/Domain: Thermotoga xylanase A amino-terminal repeat homology <TXX>

Query Match          4.6%; Score 109; DB 2; Length 800;
Best Local Similarity 19.7%; Pred. No. 0.96;
Matches 91; Conservative 51; Mismatches 150; Indels 170; Gaps 24;

Oy      30  PYAFVDEVRSKND-----LGDNDLLICVQSATQASPTDANPL-DEHEPELY-----TT 78
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      25  PTALAEGRTRDNKFKHLGNDNVKRPSEAGALQLQFVDGOMTLVDYDGEVITQLRGSMTH 84
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy      79  ALENTMMLNCASL-----NODIMRLACY-DTLVHGEPFAVITKRSIRLDETI---- 126
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

OY 127 ----WQI-----IKGRPVYVYQETDP-----IFLMDN 150
Db 145 YVIYDMHNAHARPDPDYAGAEDEFERDIAALYPPNPHIILELANEPSSNNNGAGI.PNN 204
OY 151 EKGMLTKKDAKOLEAAKQFTPLSLSPDLDRN---NPLWSSRP-----HNPMY 196
Db 205 EEGMNAVK-----EYADRIYEMLRDSGNADNIIIVGSPNMSORPDLAANDRIDHHTMY 259
OY 197 VLPIF--MHGRPNNS--PNTPSHEARQFTPNEFRAPELKFOVSYKVAALDLMGT----- 247
Db 260 TVHFYTGSHAASTESYEPREPENSEKGNVMSMTRYA---LENGAIVFATE--WCTSOANG 313
OY 248 -----DSDLMFGTQO-----SHWOIENGKNS-----REFRYH----- 275
Db 314 DGGPTDEADWDIEFLANNENNISMAWMSLTN--KNEYSGAFITPEELCKSNATSLDGPQDQYW 372
OY 276 -----DYQPELFIQTQPYVSDLPW---DGKVRMTGMGAHVHNSGES- 312
Db 373 VPPEELSLSGEYVRARIKGVNPEP---IDRTYTTVLMQFNGTQKOGGV-----NDSP 423

```

OY 313 -----AKSRSMNRAVIMAGMEKNTLTPRIMGR 342
Db 424 VEDVVIENAGALKLGLSDASNDVSEGNWYANRLSADGWCK 465

RESULT 13

UT0611
cellulase (EC 3.2.1.4), alkaline - Bacillus sp. (strain KSM-64)
N:Alternate names: endo-1,4-beta-glucanase
C:Species: Bacillus sp.
C:Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 10-Dec-1999
C:Accession: J06011
R:Sumitomo, N.; Ozaki, K.; Kawai, S.; Ito, S.
Biosci. Biotechnol. Biochem. 56, 872-877, 1992
A:Title: Nucleotide sequence of the gene for an alkaline endoglucanase from an alkalophilic
A:Reference number: J06011; MUID:9205459
A:Accession: J06011
A:Molecule type: DNA
A:Residues: 1-822 <SUM>
A:Cross-references: GB:M84963; NID:9289264; PIDN:AAA73189.1; PID:9289266

C:Function:
A:Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as ce
A:Pathway: cellulose degradation
C:Superfamily: Bacillus sp. alkaline cellulase; Thermotoga xylinase A amino-terminal re
C:Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation
F:585-726/Domain: Thermotoga xylinase A amino-terminal repeat homology <TXA>

Query Match 4.6%; Score 109; DB 2; Length 822;
Best Local Similarity 19.7%; Pred. No. 1;

Matches 91; Conservative 51; Mismatches 150; Indels 170; Gaps 24;

OY 30 PVAFVDEVRKND-----IGQDNELLIGVOSATQASDTDPANPL-DEHEPELY-----TT 78
Db 25 PTLAEGNTRKRENFHILGNDVNRKPESEAGALQLEVDQGMPLVQHGEKIKLRGKSTH 84
OY 79 ALENKMLINCAL-----NODIMRLACY-DLVHGETPRAVIKTKRSIRLDETT----- 126
Db 85 GLQWPEELINDNAYKALANDWESNMIRLAMYVGENGYASNPBLIKSKVITKIGLAIENDM 144
OY 127 -----MGT-----IKGRQVYVOETPR-----IFLMGN 150
Db 145 YVLVDHVNARCPDRDPAVYAGAEFFRDIALYPRNNPHITIELANEPSSNNGAGAPNN 204
OY 151 EKGMILTKDAKQLEVAKQFTPLISFDLDRN-----NTPLMSSRP-----HNPMY 196
Db 205 EEEGMNWK-----EYADPIYEMLRDSGNADNIIIVGSPWMSQRPDLAANPIDDHHTWY 259
OY 197 VLPIF--MGCKPNRS--PNTPSHEARQFTPRNEFRAPBLKQVSVYKKAEDLNGT----- 247
Db 260 TVHFYGSNHAASYEPPEPNSERGVMSNTRYA-----LENGVAVFATE--WGTSQANG 313
OY 248 -----DSDLMFGYTOQ-----SHMOIFNGKNS-----RPRVH----- 275
Db 314 DGGPIYDEADQVWLEFLNENNISWANNSLTN-KNEVSGAPTFEELGKSNATSLDPEPDQV 372
OY 276 -----DYQPEIFLQPVYSDLPV-----DGKVRMIGMGAVHNSGES- 312
Db 373 VPPELSLSGEYVARIKGVYEP---IDRTKYKVLMDPFNDGTRKQGGV-----NGDSP 423
OY 313 -----AKSRSMNRAVIMAGMEKNTLTPRIMGR 342
Db 424 VEDVVIENAGALKLGLSDASNDVSEGNWYANRLSADGWCK 465

RESULT 14

TVKTRR
protein kinase (EC 2.7.1.37) raf - rat
N:Alternate names: kinase-related transforming protein raf; raf proto-oncogene protein-s
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 11-Jun-1999
C:Accession: B26126

R: Ishikawa, F.; Takaku, F.; Nagao, M.; Sugimura, T.
Mol. Cell. Biol. 7, 1226-1232, 1987
A:Title: Rat c-raf oncogene activation by a rearrangement that produces a fused prote
A:Reference number: A26126; MUID:87172791
A:Accession: B26126
A:Molecule type: mRNA

A:Residues: 1-602 <TSH>
A:Cross-references: GB:M15428; NID:9206546; PIDN:AAA42002.1; PID:9206547
A:Genes: raf

C:Superfamily: rat protein kinase raf; protein kinase homology
C:Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; serine/thre
F:501-567/Domain: protein kinase homology <KIN>
F:309-311/Region: protein kinase ATP-binding motif
F:329/Active site: Lys #status predicted
F:453/Binding site: phosphate (Ser) (covalent) (by autophosphorylation) #status predi

Query Match 4.5%; Score 106.5; DB 1; Length 602;
Best Local Similarity 19.0%; Pred. No. 1;

Matches 92; Conservative 74; Mismatches 172; Indels 145; Gaps 24;

OY 23 QAQAVRPAFVDEVRKNDLIGQDNELLIGVOSATQASDTDPANPLDEHEPELYTTALEN 82
Db 81 ELEKLNQVAVLEKKELETAQDRNL-GIQQFTRAKE-----LEAKRDLIRT---N 131
OY 83 KTMILNCASALNODIMRLACYDLVHGETPRAVIKTKRSIRLDETTWO--TIKGRQVYVOE 140
Db 132 ERLSQEVEYLTEDEVKRL--NEKLKESNT--TKGEIQLKIDELQASDVVYKREKRLQE 186
OY 141 -----TTPPIFLMGNEKG-----MLTKKDAKQLEVAKQFTPLS 174
Db 187 KELLHQNQSLNLTETKTKDELALGREGNEITELKLTLENKEDDARSSESASPSA 246
OY 175 LSFDDLNRNTPPLMSSRPHNPMYVLPFPHGKPNRSPNTPSHEARQFTPRNEFRAPBLKQV 234
Db 247 LS--SSPNLSPTGSGQKPT-----VPAQREARASGQGEKMKIRPRQGRSSYWEI 298
OY 235 SVKVAKAEDMGTDSDLMFGYTOQSHW-----QIFNGKSRPRVHDXOPEI----- 281
Db 299 EASEVMLSTRIGSGS--FGTVYKKGWHDVAVKILKLVDPRTPEQLQAFREVAVLKTR 355
OY 282 -----FLTO-----PVYSDL--PMDGKVMIT-----GMAVNH 307
Db 356 HVNILLFMGYTKDNLAIVQWCEGSSLYKHLHVOETKFMQFLDIARQTAGOMDYLAH 415
OY 308 SNGESAKLSRSMNRAVIMAGMEK-----NLTPRIMGRIFKGGSGS----- 351
Db 416 KNIIRHDKS--NNIFLHEGLTVKIGDGLATYKSR-----SSGQVQEPRTGSVLYM 466
OY 352 -----PDDNPDL--DYVGYDVRFLYQLEKNSISGTVRVNPBSGKALQLDVYVPL 402
Db 467 APEYIRQDNNPFSFGSDVSYGIV--LYEL-----MTGELPSHINNROI----IFWV 515
OY 403 GKG 405
Db 516 GRG 518

RESULT 15

JCS467
cellulase (EC 3.2.1.4) - Bacillus sp.
N:Alternate names: Endo-1,4-beta-glucanase
C:Species: Bacillus sp.
C:Date: 07-Jul-1997 #sequence_revision 18-Jul-1997 #text_change 10-Dec-1999
C:Accession: JCS467
R: Miyatake, M.; Imada, K.
Biosci. Biotechnol. Biochem. 61, 362-364, 1997
A:Title: A gene encoding endo-1,4-beta-glucanase from Bacillus sp. 22-28.
A:Reference number: JCS467; MUID:97212038
A:Accession: JCS467
A:Molecule type: DNA
A:Residues: 1-783 <MTY>

A:Cross-references: DDBJ:D85236

A:Experimental source: strain 22-28

A:Note: neither the complete nucleic acid sequence nor the complete translation are shown
C:Superfamily: Bacillus sp. alkaline cellulase; Thermotoga xylosanase A amino-terminal ref

C:Keywords: glycosidase; hydrolase
F:587-728/Domain: Thermotoga xylosanase A amino-terminal repeat homology <TXA>

```
Query Match          4.3%; Score 101.5; DB 2; Length 783;
Best Local Similarity 19.7%; Pred. No. 3.9;
Matches 91; Conservative 54; Mismatches 148; Indels 169; Gaps 26;

QY 30 PYAFVDEYRSKND-----LGQDNELLIGVQATQSSTDITANPL-DEHEPELY-----TT 78
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 25 PTALAEENGTREDNPDHLLGNENVRPSEAGALQLEVDGQMTLVDOHGEEKIQLRGMSTH 84
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 79 ALENTMLINCSAL-----NODIMRLACYDTLVHGE-----TPAVIKTK-----RSI 120
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 85 GLQWPEELINDNAYKALSNDDMSNMIRLAMY-----VGENGATNPBELIKQVYIDGIELAI 140
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 121 RLDETI---WQT-----IKGPQVYVYQETTD-----TF 146
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 141 ENDMTVIYDWHVHAPRDPDPYAGAEDEFFRDIALYPMNPHIYELANEPSSNNNGAG 200
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 147 IMGNEKGLTKKDAKQLEYAAKQFTPLSLSPDLRN-----NTPLMSSRP-----H 192
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 201 IPNNEEGMKAVK-----EXADPIVEMLRDSGNADNIIIVGSPNMSQRPDLADNPINDH 255
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 193 NPMVYLPIF--MHGKPNRS--PNTPSHEARQFTPNFRAPELKFOVSVYKKAEDLMGT- 247
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 256 HTMYTVHFTYTGSHASTESYPPEPTNSEKGNVMSNTRYA-----LENGAVVFATE--WGTS 309
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 248 -----DSDLMFQYTDQ-----SHWQIFNGKNS-----RFRV----- 274
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 310 QANGDGPYFDEADVIEFLNENNISMWMSLTN--KNEVSGAFPPFELGKSNATSLDPCP 368
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 275 -----HDYQPELFLQPYVSDLPW---DGKVRMIGMGA----- 304
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 369 DQVMAPEELSLSGEYVRAIRIKGAKYEP---IDRTRYTKVLMDFNDGTGKGFVNSDSPNK 425
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 305 ----VHHSNGESAKLSRSMNRAYLMAGMEKKNLTVMPRIWGR 342
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 426 EAIEYVENENG-TLRISGLNVSNLSDGNFMANFRLSANGMGR 466
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: November 30, 2001, 14:18:34
Job time: 256 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:27:00 ; Search time 28.02 Seconds

(without alignments)
578.367 Million cell updates/sec

Title: US-09-787-083-4
Perfect score: 2360
Sequence: 1 MKVSLSTLTSLTLPCEFLA.....YNHEATSGVGLMDNMGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues
Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	373	15.8	286	1 P37446	1 P37446 KLEPN
2	358	15.2	289	1 P37442	1 P37442 KLEPN
3	357	15.1	289	1 P37442	1 P37442 KLEPN
4	321.5	13.6	289	1 P37442	1 P37442 KLEPN
5	109	4.6	800	1 P37442	1 P37442 KLEPN
6	94.5	4.0	397	1 P37442	1 P37442 KLEPN
7	93	3.9	4273	1 P37442	1 P37442 KLEPN
8	92.5	3.9	849	1 P37442	1 P37442 KLEPN
9	92	3.9	324	1 P37442	1 P37442 KLEPN
10	92	3.9	696	1 P37442	1 P37442 KLEPN
11	92	3.9	888	1 P37442	1 P37442 KLEPN
12	91.5	3.9	969	1 P37442	1 P37442 KLEPN
13	91	3.9	496	1 P37442	1 P37442 KLEPN
14	91	3.9	1474	1 P37442	1 P37442 KLEPN
15	90.5	3.8	467	1 P37442	1 P37442 KLEPN
16	90	3.8	1158	1 P37442	1 P37442 KLEPN
17	90	3.8	1788	1 P37442	1 P37442 KLEPN
18	89.5	3.8	486	1 P37442	1 P37442 KLEPN
19	89.5	3.8	669	1 P37442	1 P37442 KLEPN
20	89.5	3.8	1478	1 P37442	1 P37442 KLEPN
21	89.5	3.8	1840	1 P37442	1 P37442 KLEPN
22	89	3.8	1426	1 P37442	1 P37442 KLEPN
23	88.5	3.8	825	1 P37442	1 P37442 KLEPN
24	88.5	3.8	992	1 P37442	1 P37442 KLEPN
25	88	3.7	810	1 P37442	1 P37442 KLEPN
26	88	3.7	1115	1 P37442	1 P37442 KLEPN
27	88	3.7	1382	1 P37442	1 P37442 KLEPN
28	87.5	3.7	844	1 P37442	1 P37442 KLEPN
29	87.5	3.7	1379	1 P37442	1 P37442 KLEPN
30	87	3.7	503	1 P37442	1 P37442 KLEPN
31	87	3.7	666	1 P37442	1 P37442 KLEPN
32	87	3.7	842	1 P37442	1 P37442 KLEPN
33	87	3.7	882	1 P37442	1 P37442 KLEPN

ALIGNMENTS

34	86	3.6	882	1	HSS1_HUMAN	P52848	homo sapien
35	86	3.6	1087	1	XYNX_CLOTM	P38535	clostridium
36	85.5	3.6	353	1	DCUP_BACSV	P22395	bacillus su
37	85.5	3.6	790	1	SEIL_MOUSE	Q92426	mus muscicu
38	85.5	3.6	794	1	SEIL_HUMAN	Q9ubv2	homo sapien
39	85	3.6	403	1	P37_MYCHR	P15363	mycoplasma
40	85	3.6	491	1	TY3H_PHASP	P11982	phasiandae
41	85	3.6	560	1	DTXH_CORBE	P00589	corynebacte
42	85	3.6	828	1	BGAL_BRAOL	P49676	brassica ol
43	85	3.6	1184	1	ALAC_ARATH	P57792	arabidopsi
44	85	3.6	2231	1	SENI_YEAST	Q00416	saccharomyc
45	84.5	3.6	478	1	DHGB_ACICA	P13650	actinobact

RESULT 1	P37446	STANDARD	PRT	286 AA.
ID	P37446			
AC	P37446			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT			
DE	PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-			
DE	ACTYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).			
GN	PLDA.			
OS	Klebsiella pneumoniae.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Klebsiella.			
OX	NCBI_TaxID=573;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94131966; PubMed=8300539;			
RA	Brok R.G.P.M., Brinkman E., van Boxtel R., Bekkers A.C.A.P.,			
RA	Verheij H.M., Tomassen J.;			
RT	"Molecular characterization of enterobacterial plda genes encoding			
RT	outer membrane phospholipase A.";			
RL	J. Bacteriol. 176:861-870(1994).			
CC	-1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE			
CC	A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.			
CC	-1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O -> 1-ACYLGLYCERO-			
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.			
CC	-1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O -> 2-ACYLGLYCERO-			
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.			
CC	-1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.			
CC	-1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES			
CC	LOCATED THERE.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL: X76901; CA54223.1; -			
DR	PIR: B36971.			
DR	PIR: S40129; S40129.			
DR	InterPro: IPR003187; PLA1.			
DR	Pham; PF02253; PLA1; 1.			
KW	Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium.			
FT	SIGNAL	1	20	BY SIMILARITY.
FT	CHAIN	21	286	PHOSPHOLIPASE A1.
FT	ACT_SITE	161	161	BY SIMILARITY
SO	SEQUENCE	286 AA;	3E39F863085108A3	CRC64;

Query Match	15.8%;	Score 373;	DB 1;	Length 286;
Best Local Similarity	42.1%;	Pred. No. 3.3e-23;		
Matches	90;	Conservative	27;	Mismatches 87; Indels 10; Gaps 5;

Query Match 15.2%, Score 358, DB 1, Length 289;

Best Local Similarity 39.3%; Pred. No.5.6e-22;
Matches 84; Conservative 30; Mismatches 90; Indels 10; Gaps 4;

```
QY 226 RAPELKEFVYSKVKRAAEIDMGTDSDLMGTYQDSHWQIFPNCKNSRPFVHYDQPEIFLTQ 285
      +:::|::|: |::|: |::|: |::|: |::|: |::|: |::|: |::|: |::|: |::|: |::|:
Db 83 RKDEKFKQLSLAFPLMRGILGPNVSLGASYSOKSWMLSNSKSSSPFRFTNEPOLFLGF 142
QY 286 PV-YSDLPWDCKVRNIGGAVHNSGESAKLSRSNNRAYLMAGMKMLTVPRIMGRIE 344
      +::|::|: |::|: |::|: |::|: |::|: |::|: |::|: |::|: |::|: |::|: |::|:
Db 143 ATDYRFNAQ--TLRIVEKGYNHDSNGRDPDPSRTNNRLYTLRLMAENGKMLVEYKRWYVI- 199
QY 345 KEGSGSPDDNDPDLIDYGYGDVRFLYOLENKSNTISGTVRYNPRSGKALDLDYVPLGK 404
      ||||| |::|: |::|: |::|: |::|: |::|: |::|: |::|: |::|: |::|: |::|:
Db 200 -----GSTDDNDPDLIRKGYGYOLKIGYHL-GEAVYLSAGQYNNMTGYGGAEGVGLSPYTK 253
QY 405 GISGFOIEFGYGGSLIDYNHEATSFEGMLMND 438
Db 254 HVLITGVYSGTGESLIDYNFNQTRKVGVMLEND 287

RESULT 3
PAL_ECOLI STANDARD: PRT: 289 AA.
AC P00631:
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT
DE PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYCHOLINE 1-
DE ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).
GN PLDA OR B3821 OR Z5342 OR ECS4751.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_Taxid=562, 83334;
[1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE=85157492; PubMed=6397464;
RA Homma H., Kobayashi T., Chiba N., Karasawa K., Mizushima H., Kudo I.,
RA Inoue K., Ikeda H., Sekiguchi M., Nojima S.;
RA "The DNA sequence encoding plda gene, the structural gene for
RA detergent-resistant phospholipase A of E. coli.";
RA J. Biochem. 96:1655-1664(1984).
[2]
RN RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RC STRAIN=K12 / MG1655;
RA Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;
RA "Analysis of the Escherichia coli genome: DNA sequence of the region
RA from 84.5 to 86.5 minutes.";
RA Science 257:771-778(1992).
[3]
RN RP REVISION TO 14-15.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12.";
RA Science 277:1453-1474(1997).
[4]
RN RP STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobbeck E.J., Davis N.W., Llm A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
```

RL Nature 409:529-533(2001).
 RM [5]
 RP SEQUENCE FROM N.A.
 RX STRAIN=O157:H7 / RIMD 0509952;
 RC MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kunita S., Shiba T., Hattori M., Shinagawa H.;
 RA "Complete genome sequence of enterohemorrhagic *Escherichia coli*
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 RM [6]
 RP SEQUENCE OF 30-289 FROM N.A.
 RX STRAIN=K12;
 RC MEDLINE=85003590; PubMed=6383820;
 RA de Gans P., Verheij H.M., Riegman N.H., Hoekstra W.P.M., de Haas G.H.;
 RA "The pro- and mature forms of the *E. coli* K-12 outer membrane
 RT phospholipase A are identical.";
 RL EMBO J. 3:11799-1802(1984).
 RM [7]
 RP SEQUENCE OF 174-289 FROM N.A.
 RX STRAIN=K12;
 RC MEDLINE=87115164; PubMed=3027506;
 RA Irino N., Nakayama K., Nakayama H.;
 RA "The *recQ* gene of *Escherichia coli* K12: primary structure and
 RT evidence for SOS regulation.";
 RL Mol. Gen. Genet. 205:298-304(1986).
 RM [8]
 RP MUTAGENESIS OF SFR-172.
 RX MEDLINE=94131966; PubMed=8300539;
 RA Bick R.G.P.M., Brinkman E., van Bontel R., Bekkers A.C.A.P.,
 RA Verheij H.M., Tommassen J.;
 RA "Molecular characterization of enterobacterial *pilA* genes encoding
 RT outer membrane phospholipase A.";
 RL J. Bacteriol. 176:861-870(1994).
 RM [9]
 RP ACTIVE SITE SER-164, AND PARTIAL SEQUENCE.
 RX MEDLINE=91249806; PubMed=2040286;
 RA Horrovoets A.J.G., Verheij H.M., de Haas G.H.;
 RA "Inactivation of *Escherichia coli* outer-membrane phospholipase A by
 RT the affinity label hexadecanesulfonyl fluoride. Evidence for an
 RL active-site serine.";
 RL Eur. J. Biochem. 198:247-253(1991).
 CC -1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE
 CC A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
 CC REQUIRED FOR EFFICIENT SECRETION OF BACTERIOCINS, SEEMS TO BE
 CC DORMANT IN NORMAL GROWING CELLS.
 CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-
 CC PHOSPHOCHOLINE + A FATTY ACID ANION.
 CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-
 CC PHOSPHOCHOLINE + A FATTY ACID ANION.
 CC -1- COPACOPOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
 CC LOCATED THERE.
 CC -1- INDUCTION: BY MEMBRANE DAMAGING, FOR EXAMPLE, BY PHAGE-INDUCED
 CC LYSIS OR TEMPERATURE SHOCK.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X02143; CAA26081.1; -
 DR EMBL: M87049; AAA67617.1; -
 DR EMBL: AE000458; AAC76824.1; -
 DR EMBL: AE005613; AAC59017.1; -
 DR EMBL: AP002567; BAB38174.1; -
 DR EMBL: M30198; AAA24516.1; -
 DR PIR: A00771; PSECA.

```

DR      PIR: A22133; PSECA1.
DR      PIR: S30711; S30711.
DR      Ecogene; EG10738; plda.
DR      InterPro; IPR003187; PLAL.
DR      Pfam; PF02253; PLAL; 1.
KW      Hydroxylase; lipid degradation; Outer membrane; Signal; Calcium; Complete proteome.
FT      SIGNAL             1          20
FT      CHAIN              21         289    PHOSPHOLIPASE A1.
FT      ACT_SITE           164        164
FT      MUTAGEN            172        172    S->F: INACTIVE PROTEIN.
FT      CONFLICT           14         15    LP -> FA (IN REF. 2.).
FT      CONFLICT           30         33    DAPA -> MTRQ (IN REF. 6).
SQ      SEQUENCE          289 AA; 33163 MW; A688AD32AA60F218 CRC64;

Query Match               15.1%; Score 357; DB 1; Length 289;
Best Local Similarity     39.3%; Pred. No. 6,7e+22;
Matches   84; Conservative 29; Mismatches 91; Indels 10; Gaps 4;

OY      226 RAPELKFQVSVKVKAAEDLMGTDSLMFGYYTQQSHMOJFNCKNSRPFVHDYOPEIFLTQ 285
DB      83 RDEVEKFKQLSLAFLPLMRGILIGPNSVLGASYTQKSMWOLSNBESSPFRETNEPEOLFGE 142
OY      286 PV-YGDLPWDKGAVRRATGCAVHNSGESAKLSRSNNRYLAAGMKMKTLYTPRTWGRIIF 344
DB      143 ATDYEPFAGV-"TLRDVENGMYNHDSNGSRDPTGRSNNRLYTRIAMAENGMMLEVKPMYYV- 199
OY      345 KRGSSGSOPDDNDLLIDYGYGVDFRFXOLENKSNTSGTVRPNRSGKALOLDYVPGLK 404
DB      200 ----GNTDDNDPDITKYMGTYQLKGTHL-GDAVLSAKGQIWMNTYGGAELGLSTPYTK 253
OY      405 GISGFYLFPGQGSGSLIDYNHEATSFEGVGLMND 438
DB      254 HYLRTQYYSGGESGLIDINFQTRVGVMIND 287

RESULT  4
PAL_PROVU ID PAL_PROVU STANDARD; PRF; 289 AA.
AC P37447;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETURGENT-RESISTANT
DE PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPATTIDYLCOLINE 1-
DE ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).
GN PLDA.
OS Proteus vulgaris.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Proteus.
OX NCBI_TaxID=585;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94131966; PubMed=8300539;
RA Brok R.G.P.M., Brinkman E., van Boxtel R., Bekkers A.C.A.P.,
RA Verheij H.M., Tommassen J.;
RT "Molecular characterization of enterobacterial plda genes encoding
RT outer membrane phospholipase A."
RL J. Bacteriol. 176:861-870(1994).
CC -1- FUNCTION: HYDROLYSIS OF PHOSPATTIDYLCOLINE WITH PHOSPHOLIPASE
CC A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
CC -1- CATALYTIC ACTIVITY: PHOSPATTIDYLCOLINE + H(2)O = 1-ACYLGYCERO-
CC PHOSPHOCOLINE + A FATTY ACID ANION.
CC -1- CATALYTIC ACTIVITY: PHOSPATTIDYLCOLINE + H(2)O = 2-ACYLGYCERO-
CC PHOSPHOCOLINE + A FATTY ACID ANION.
CC -1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
CC LOCATED THERE.
```

CC use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/> or send an email to license@isb-slb.ch).

*Molecular cloning, expression, and characterization of
RT endo-beta-1,4-glucanase genes from *Bacillus polymyxa* and *Bacillus*
RT *circulans*.*
RL J. Bacteriol. 172:1576-1586(1990).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
CC LINKAGES IN CELLULOSE.
CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY A (FAMILY 5 OF GLYCOSYL
CC HYDROLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M33791; AAA22631.1; -
DR PIR: A35136; A35136.
DR HSSP: P54583; 1ECE.
DR InterPro: IPR001547; Glyco_hydro_F5.
DR Pfam: PF00150; cellulase_1.
DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; 1.
KW Cellulose degradation; Hydrolase; Glycosidase.
FT ACT_SITE 194 194 PROTON DONOR (BY SIMILARITY).
FT ACT_SITE 317 317 NUCLEOPHILE (BY SIMILARITY).
SQ SEQUENCE 397 AA; 44357 MW; B9C2E802C04F0A2A CRC64;

Query Match 4.0%; Score 94.5; DB 1; Length 397;
Best Local Similarity 25.7%; Pred. No. 2.7;
Matches 46; Conservative 21; Mismatches 61; Indels 51; Gaps 10;

QY 301 GCGAVHNSGSAKLSRNRAYLAGMKNLTVMRIGR-----IFKEG----- 347
DB 43 GNRIVDSGKEAFNGLNW-----FGLERPNTY-LHGLMSRSDMDLDQVKEGYNLIR 95
QY 348 -----SGSQPDD-----NPDIIIDYGVGVRFYOLENKSNSIGTV-----RYNPRS 389
DB 96 LPYSNOLFDSRRPDSIDYKKNPDLY---GLNPQIMDKLIERAGGQIITIDRRHPPGS 152
QY 390 GKGALDLDYV--YPLGKISGVFOIFQGY-----GSLIDYHNEATSPGVGLMLNDW 439
DB 153 G-CQSELMWTISQYPPESRWMISDMKMLADRYKNNPVIYGADLHNEPHGASNGTGNASTDW 210

RESULT 7
PKSM_BACSU STANDARD; PRT: 4273 AA.
AC P40872: O31781;
DT 01-FEB-1995 (Rel. 31, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PUTATIVE POLYKETIDE SYNTHASE PKSM.
GN PKSM OR PKSY.
OS *Bacillus subtilis*.
OC Bacteria; Firmicutes; *Bacillus*/Clostridium group;
OC *Bacillus*/Scaphylococcus group; *Bacillus*.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-1763 FROM N.A.
RC STRAIN=168 / Pbl1424;
RA Tognoni A., Grandi G.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: POTENTIALLY INVOLVED IN SOME INTERMEDIATE STEPS FOR
CC THE SYNTHESIS OF A POLYKETIDE MOLECULE WHICH MAY BE INVOLVED IN
CC SECONDARY METABOLISM.
CC -1- COFACTOR: CONTAINS 4 COVALENTLY BOUND PHOSPHOPANTETHEINES

(POTENTIAL).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z99113; CAB3603.1; -
DR EMBL: Z35133; CAA84505.1; -
DR Subtilist; BG10931; PKSM.
DR InterPro: IPR000794; ketoacyl-synt.
DR InterPro: IPR001601; Meth-transf.
DR InterPro: IPR003880; Phosphopant_attach.
DR Pfam: PF00109; ketoacyl-synt; 1.
DR Pfam: PF00550; pp-binding; 1.
DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 2.
DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 2.
DR PROSITE: PS50075; ACP_DOMAIN; 4.
KW Transferase; Acyltransferase; Antibiotic biosynthesis; NADP;
FT Phosphopantetheine; Multifunctional enzyme; Repeat; Complete proteome.
FT DOMAIN 295 364 ACRYL CARRIER (ACP) 1.
FT DOMAIN 396 834 BETA-KETOACYL SYNTHASE 1.
FT DOMAIN 2190 2238 ACRYL CARRIER (ACP) 2.
FT DOMAIN 2322 2737 BETA-KETOACYL SYNTHASE 2.
FT DOMAIN 3532 3947 BETA-KETOACYL SYNTHASE 3.
FT DOMAIN 3410 3483 ACRYL CARRIER (ACP) 3.
FT DOMAIN 4140 4209 ACRYL CARRIER (ACP) 4.
FT BINDING 327 327 PHOSPHOPANTETHEINE (POTENTIAL).
FT BINDING 2222 2222 PHOSPHOPANTETHEINE (POTENTIAL).
FT ACT_SITE 2476 2476 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
FT ACT_SITE 3446 3446 PHOSPHOPANTETHEINE (POTENTIAL).
FT ACT_SITE 3690 3690 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
FT BINDING 4172 4172 PHOSPHOPANTETHEINE (POTENTIAL).
FT CONFLICT 103 103 E -> V (IN REF. 2).
FT CONFLICT 276 276 Q -> E (IN REF. 2).
FT CONFLICT 289 289 T -> S (IN REF. 2).
SQ SEQUENCE 4273 AA; 477459 MW; 3BBFC1A250AEB5A CRC64;

Query Match 3.9%; Score 93; DB 1; Length 4273;
Best Local Similarity 19.2%; Pred. No. 97;
Matches 102; Conservative 79; Mismatches 163; Indels 188; Gaps 29;

QY 14 PCFALIAIOOAAV-PMPVAFVDE-----VRSKN-----DLGQ 45
DB 3597 PLFQIISPKRESMDPROPIFLEAMHTFEDAGYMGRIKSGCYVVGVEGYAHLTG 3656
QY 46 DNEILLIGVOSATOSTDTANPLDHEPELEYTALENKTMILI---NCSALNQDITRLACY 102
DB 3657 DTDYINGTQATLSAR--IAVALDKGPNNALTAACSSGLVAIHQACSLAROG---DCE 3710
QY 103 DTLVHGEPRAVITKRSIRLDETITWOTIKGRPOVVYQETDPTILMKNENKMLTKKAKO 162
DB 3711 MALAAGVTLN-----SHMSFALTRAEMLSPNGCKYFQDQANG 3750
QY 163 L-----EYAAKQFTPLSLFDLRNNTPLMSGRPNPMVYLIFMHGKPN--RSPNTSHE 216
DB 3751 LVREAVAAVILKPLSKAIE-DKDH-----YGLKASGVNVDGNTNGTIANP-PFSQ 3801
QY 217 ARO-----FTPNEFRAPLKFQVS-----VKRAAEDLWGTSDLMFGYTOQSHWQ 262
DB 3802 ALELTENIYERKEINPLDIQYVMAHSTGSLNDPLEVQALTSVFSK-----YTKROFC 3854
QY 263 IFNCKNRPRVHYQPEIFLTQPVYSDLPWDGKVRHIGM-----GAVHNSGES- 312
DB 3855 MIS--SIKPLIGHTEA-----SGTVALISMAMKNOIIPATHHCESENP 3898
QY 313 -----AKLSRSMNRAYLAGMKNLTVMRIGRIFKEG--SGSQ----- 351

Db 3699 YIPKESFVLCKENSRWK-----KN--OKPRN-GTISTGTGNTAHAVIEYI 3946
 QY 352 PDDPDLIDYGYGV-----RFLYLENKSNI-S-GVRYNPRSGCAL 394
 Db 3947 PDDPSTORHOGSFOIFPISANDRLDADACRMATVLEONHNSLPRVATTLQVGRAM 4006
 QY 395 QL-----DYVPL-GKISGYEQ-IFQYQGLIDYNEHA 427
 Db 4007 EARLAIYVNOEQVLKRYLKEYVEAMKNGVSGQGLYGTGTEGLEBODEA 4058
 RESULT 8
 SRK6_BRAOL STANDARD; PRT: 849 AA.
 AC Q009092;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE PUTATIVE SERINE/THREONINE KINASE RECEPTOR PRECURSOR (EC 2.7.1.37)
 DE (S-RECEPTOR KINASE) (SRK).
 GN SRK6.
 OS Brassica oleracea (Cauliflower).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Brassica.
 OC NCBI_TaxID=3712;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPRAIN-CV_S6S6: TISSUE=Stigma;
 RX MEDLINE=92020942; PubMed=1681543;
 RA Stein J.C., Howlett B., Boyes D.C., Nasrallah M.E.;
 RT "Molecular cloning of a putative receptor protein kinase gene encoded
 RL at the self-incompatibility locus of Brassica oleracea.";
 Proc. Natl. Acad. Sci. U.S.A. 88:8816-8820(1991).
 CC -1- FUNCTION: INVOLVED IN SPOROPHYTIC SELF-INCOMPATIBILITY SYSTEM
 (THE INABILITY OF FLOWERING PLANTS TO ACHIEVE SELF-
 CC FERTILIZATION), PROBABLY ACTING IN COMBINATION WITH S-LOCUS-
 CC SPECIFIC GLYCOPROTEINS. INTERACTION WITH A LIGAND IN THE
 CC EXTRACELLULAR DOMAIN TRIGGERS THE PROTEIN KINASE ACTIVITY OF THE
 CC CYTOSOLIC DOMAIN.
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN -> ADP + A PHOSPHOPROTEIN.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY IN THE PISTIL AND ANTHER.
 CC -1- POLYMORPHISM: THERE ARE A NUMBER OF DIFFERENT S ALLELES IN
 CC B. OLERACEA, POSSIBLY PROVIDING THE RECOGNITION SPECIFICITY.
 CC -1- SIMILARITY: THE EXTRACELLULAR DOMAIN IS SIMILAR TO S-LOCUS
 CC GLYCOPROTEINS OF BRASSICA, WHILE THE INTRACELLULAR DOMAIN IS
 CC A SER/THR-PROTEIN KINASE RELATED TO RAF KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M76647; AAA33000.1; ALT_TERM.
 DR HSSP: P1362; JGI.
 DR InterPro: IPR001480; B_lectin.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR003609; Pan_app.
 DR InterPro: IPR002290; Ser_thr_kin_actsite.
 DR InterPro: IPR000858; Slocus_glycop.
 DR Pfam: PF00069; pkinase.1.
 DR Pfam: PF00954; S_locus_glycop.1.
 DR SMART: SM00473; B_lectin.1.
 DR SMART: SM00473; PAN_AP.1.
 DR SMART: SM00221; STYC.1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST.1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM.1.
 KW transferase; Serine/threonine-protein kinase; signal; ATP-binding;

KW Transmembrane; Receptor; Glycoprotein; Self-incompatibility.
 FT SIGNAL 1 32
 FT PUTATIVE SERINE/THREONINE KINASE
 FT CHAIN 33 849
 FT
 FT DOMAIN 33 446
 FT TRANSMEM 447 466
 FT DOMAIN 467 849
 FT DOMAIN 528 779
 FT NP_BIND 534 542
 FT BINDING 556 556
 FT ACT_SITE 653 653
 FT CARBOHYD 47 47
 FT CARBOHYD 120 120
 FT CARBOHYD 196 196
 FT CARBOHYD 260 260
 FT CARBOHYD 314 314
 FT CARBOHYD 389 389
 FT CARBOHYD 442 442
 SQ SEQUENCE 849 AA; 97231 MW; 7E156059EDDF4370 CRC64;
 Query Match 3.9%; Score 92.5; DB 1; Length 849;
 Best Local Similarity 21.9%; Pred. No. 11;
 Matches 78; Conservative 50; Mismatches 107; Indels 121; Gaps 24;
 QY 105 LVHGETPAVVKTKRSIRLDETITQIKGPOVYQETDPIFLMG-----NEK--GMLTK 157
 Db 24 LIHPALSTIYNT-----LSTESTLTSNMKTLV-----SPQSIREFGVFFRNSRWYLGWMYK 75
 QY 158 K-DAKOLEYAKOFTPLSLSPD-----LDRNNTPLMSSRPHNPMYLYLFFMNGK 205
 Db 76 KVSDDTYVWVNRDPLNSAIGTLKISGNLVLLDHSNKPMWMTN-----LTRGN 125
 QY 206 PRRSP-----NTPSHEAROPT-----PNEFRAPELKFGVSVKKAEDLM 245
 Db 126 -ERSVYVVELANGFNWRDSSNNDASEYLWQSFYPTDITLPEKKLGYNLKT----- 177
 QY 246 GTSDLMFGYTOQSHMOJFNGKNSRPFYVDYQ-----PELFLNQ--PYVSDLPWDGK 296
 Db 178 GLNRFL-----TSWRSSDDPSSGNF---SYKLEFQSLPEFYLSRENPMPHRSQPMWG- 226
 QY 297 VAMIGKGVH-----HNGESAKLSRSMNAY-----LMAEMKKNLTVP--RI 339
 Db 227 IFFSIPEDOKLSYVWYNEFENNEVAYTFMTNNSFSRLTLEGGYFORLTWPSIRI 286
 QY 340 WCRIRKESGSGOPDDNPDLIDY-----GYGDVRFLEYOLENKSNIQVYR-VNPRS 389
 Db 287 WNRFW-----SSPVD-PQCDTYIMCGPYAYCDY-----NTSPVCNCIQGPNPRN 329
 RESULT 9
 Y7L0_ARCFU STANDARD; PRT: 324 AA.
 ID Y7L0_ARCFU
 AC 028112;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHETICAL PROTEIN AF2170.
 GN AF2170.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
 OC Archaeoglobus.
 OC NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kerlavage A.R., Graham G., Kyrtides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirtness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,

RA Overbeek R., Gocayne J.D., Meldman J.F., McDonald L., Ufferbach T.,
 RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
 RA Sadow P.M., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon *Archaeoglobus fulgidus*.";
 RL Nature 390:364-370(1997).
 CC -1- SIMILARITY: CONTAINS 2 KELCH REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE000955; AAB89093.1; .
 DR TIGR: AF2170; .
 DR InterPro: IPR001798; Kelch.
 DR Pfam: PF01344; Kelch; 2.
 KW Hypothetical protein; Repeat; Complete proteome.
 FT REPEAT 229 276 KELCH 1.
 FT REPEAT 277 323 KELCH 2.
 SQ SEQUENCE 324 AA; 36025 MM; C445388CFEB96E45 CRC64;

Query Match 3.9%; Score 92; DB 1; Length 324;
 Best Local Similarity 20.1%; Pred. No. 3.2;
 Matches 54; Conservative 44; Mismatches 72; Indels 98; Gaps 18;

OY 224 EFRAPELKFOYSKVKAAED-----LKGTSDLMEFGYTOOSHWOLF-NGKNSRPFVH 275
 DB 63 EFPHPN-RLRI-ILLSSEDFSEORAAVWQEEILIFGT-----VFENKYSPTDQIL 112
 OY 276 DYOEPIRLTQPVYSDLP-----W-DGKVRMI-----GAGVAHNSGESAKLSRS 318
 DB 113 SFNPKLERLKNLSLPHRPSDVAANWGDNRVYIFLNNSCEYAYAFPSSESRAKLDS 172
 OY 319 -----W-NRAYL-----MAGHEMKNLVMPRIWGRIFREGSGSQ 351
 DB 173 CPHEPGCVHSHVWYGKAYFCGEGVASFDPMGFKM--IATLDYRWVAARAVADG-- 228
 OY 352 PDDNPDLIDYGYGDVAFVLQLEKNSISGT---VKNPSGKALDLYVPLGKC-- 405
 DB 229 -----YFAIGSSGSAIETKEIIRNPXTGE-LCEMRKLPVARQOA 270
 OY 406 --ISG-YFOIF-----OGYGSLIDYNH 425
 DB 271 VAVGGEYIYIFGTYTKDGYANEIIRYD 298

RESULT 10
 SCRL_SCHPO STANDARD; PRT: 696 AA.
 AC PF0530; O9UTP3;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE SRINE/THREONINE-PROTEIN KINASE SCRL (EC 2.7.1.37).
 GN SCRL OR SPAC1B9.02C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96120227; PubMed=7498728;
 RA Jin M., Fujita M., Culley B., Apollinario E., Yamamoto M.,
 RA Maundrell K., Hoffman C.;
 RT "scrl, a high copy number suppressor of defects in the CAMP-dependent

RT protein kinase pathway in fission yeast, encodes a protein homologous
 RT to the Saccharomycetes cerevisiae Sch9 kinase.";
 RL Genetics 140:457-467(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA McQuigley R.C., Rajandream M.A., Barrell B.G., Whitehead S.,
 RA Churcher C.M.;
 RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN = ADP + A PHOSPHOPROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC CAMP SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D38108; BAA07286.1; .
 DR EMBL: AL109951; CAB53053.1; .
 DR HSSP: P05132; ICTP.
 DR InterPro: IPR000008; C2.
 DR InterPro: IPR000719; Euk_Pkinase.
 DR InterPro: IPR000961; Pkinase_C.
 DR InterPro: IPR002290; Ser_Thr_kin_actsite.
 DR Pfam: PF00069; Pkinase_1.
 DR Pfam: PF00433; Pkinase_C; 2.
 DR SMART: SM00239; C2; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR SMART: SM00133; S_TKc_X; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW Transferase: Serine/threonine-protein kinase; ATP-binding; CAMP.
 FT DOMAIN 302 563 PROTEIN KINASE.
 FT NP_BIND 308 316 ATP (BY SIMILARITY).
 FT BINDING 331 331 ATP (BY SIMILARITY).
 FT ACT_SITE 428 428 BY SIMILARITY.
 FT CONFLICT 199 199 A -> R (IN REF. 1).
 SQ SEQUENCE 696 AA; 78594 MM; A7B05F5E4D42A7 CRC64;

Query Match 3.9%; Score 92; DB 1; Length 696;
 Best Local Similarity 19.2%; Pred. No. 9.3;
 Matches 87; Conservative 66; Mismatches 163; Indels 138; Gaps 19;

OY 16 FAIIAIOQAQ-VPRPVAFVD-----EVRSKN-----DLGQDNELLIG----- 52
 DB 161 YAVITFEKTYQVWPPPFDDIGGIPISIPSKNRPPLAGSAGSSGLHSELMADYRCPHW 220
 OY 53 ----VQATQASSTDTANPLDNEHEPELYTALNKTMLINCASALNODIMRLACYDTLVHG 108
 DB 221 DPEIVFDTFKKSSQMWVSYVDKEDKFLGSVKITPIPLH-----EVQENAWKL----- 270
 OY 109 ETPAVIKTKRSIRLDETIMQTIK-----GKPVVYQETDPIFLM 148
 DB 271 EPLDITKLEGEIKETIYEHENHVRGPEDEFTALRLIGKSTFGQVYLVKRNDRIRIYAM 330
 OY 149 GNEKGMILTKRAKOLEYAAKQFTPLSLSFDDDRNTPMLSSRPNHP----- 194
 DB 331 -----KTSKSLIVRKKEVT-----HTGERNIIIVRTSLSDSPFIVGLKFSQTASD 377
 OY 195 MYVLPIFMHGKPRNSPNTPSHEARQFTNPEFRAPF--LKFOVSYVKAEDLMGTDSLM 252
 DB 378 LYLTIDYMSG-----GELFNHLOHGBRFPQRAKFIYAEVLVLALEHL----- 419
 OY 253 FGYYTQSHMOIFNGKNSRPFVHVDYQPEIFLTQPVYSDLPMDGKRYMIGMAVHNSGES 312
 DB 420 -----HKHDIIY-----RDLKPENIL-----LDADGHATLDCFGISKANLSAN 457

QY 313 AKLSRSMNRAYLMA---MEMKNLTVMPIRMG---RIEKGSGSQPDNDPILDIY---G 363
 DB 458 ATTNFCGTTETLEAEVLEEDKGYTKQVDFMSIGLVFEEMCGMSPFYAPDQOMYRNA 517
 QY 364 YGDVRE---LYOLEKNSINSGVRYRNPSSGKAL 394
 DB 518 FGVKRFPGKVLSEGRSFRGLNRPNHRIGAV 551

RESULT 11

YJHO_YEAST STANDARD: PRT: 888 AA.

AC P40361;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE HYPOTHEORETICAL 104.3 KDA PROTEIN IN SMC3-KRP18 INTERGENIC REGION.
 GN YJ1070C OR J1095 OR HRD888.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 CX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C;
 RX MEDLINE=95282514; PubMed=7762302;
 RA Vandenbol M., Durand P., Dion C., Portetelle D., Hliger F.;
 RT "Sequence of a 17.1 kb DNA fragment from chromosome X of
 RT Saccharomyces cerevisiae includes the mitochondrial ribosomal protein
 RT L8.";
 RL Yeast 11:57-60(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C;
 RA Sor F.J.;

Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 -1- SIMILARITY: BELONGS TO THE ADENOSINE AND AMP DEAMINASES FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: Z34288; CAA84052.1;
 DR EMBL: Z49345; CAA89362.1;
 DR EMBL: Z49344; CAA89361.1;
 DR EMBL: X88851; CAA61309.1;
 DR PIR: S47120; S47120.
 DR SGD: S0003606; YJ1070C.
 DR InterPro: IPR001365; A.deaminase.
 DR Pfam: PF00962; A.deaminase.1.
 KM Hypothetical protein; Hydrolyase.
 SQ SEQUENCE 888 AA; 104263 MW; DB31A8086224114D CRC64;

Query Match 3.9%; Score 92; DB 1; Length 888;

Best Local Similarity 22.1%; Pred. No. 13; Mismatches 126; Indels 70; Gaps 15;

Matches 66; Conservative 36; Mismatches 126; Indels 70; Gaps 15;
 QY 128 QTIKGRPOVVOE---TTPIFLNGNEKMLTKKDA-----KOLEYAAKQFPLSL 176
 DB 155 VEAFTSTASSEAA---DTH-TEVDLKVSNANAMLSKNGIKRISIVEENMTSDIV-A 209
 QY 109 ETPAVIKTRSRJRLDEITQITKGR-POVVOETTPIFLNGNEKMLTKKDAKOLEYAA 167
 DB 210 LTFEELIKALNKWFSD---DAIKGTSTSLTKRLKDIV-----ASFLLKDSKLAVPF 259
 QY 168 KQFTPLSL-SFD-----LDRNTPLMSSRPHPM-----YLPFIHMGKPNRSP 210
 DB 260 KADTIINMPAFNVAQNTKKEIDVWDSMPVQDAKSGVSNMNGVQLVISMAGAPKNKS 319
 QY 221 TPNERRAPLTKFOVSKVKAADLMG-----TDSDLMEFYTG--QSHWQIFNKKNSRPR 273
 DB 122 FLPNPF---ELVSQMKRKYIAASKQDISINIKNDTEKWFLLPKPLPKFWREDDK---R 173

QY 274 VHDYQPELFTQPPVSDLEPMDGKVRMIGAGVAHNSGESAKLSRSMNRAYLAGHEWKNL 333
 DB 174 FQD-----PDSDLNDDGDSGTGAATPHRHGYYPPYFTDHYHYXTKSGLKGN 224
 QY 334 TVMPRIWGRIFKEGSGSQPDNDPILDIYGYGDVRFPLYOLEKNSINSGVRYRNPSSG 391
 DB 225 IKVP-YTGEYF-----DLEDY---KKQYIYHLSNQENTQNP--SPYSSK 263

RESULT 12

SACB_STRSL STANDARD: PRT: 969 AA.

AC 055242;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE LEVANSUCRASE PRECURSOR (EC 2.4.1.10) (BETA-D-FRUCTOFURANOSYL
 DE TRANSFERASE) (SUCROSE 6-FRUCTOSYL TRANSFERASE).
 GN FTF.
 OS Streptococcus salivarius.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 CC Streptococcus.
 CX NCBI_TaxID=1304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 25975;
 RX MEDLINE=9332332; PubMed=8331080;
 RA Rathnam C., Giffard P.M., Jacques N.A.;

RT "The cell-bound fructosyltransferase of Streptococcus salivarius: the
 RT carboxyl terminus specifies attachment in a Streptococcus gordonii
 RT model system.";
 RL J. Bacteriol. 175:4520-4527(1993).
 CC -1- CATALYTIC ACTIVITY: SUCROSE + (2,6-BETA-D-FRUCTOSYL)(N) -
 CC GLUCOSE + (2,6-BETA-D-FRUCTOSYL)(N+1) (OTHER SUGARS CAN
 CC ALSO ACT AS FRUCTOSYL ACCEPTORS).
 CC -1- SUBCELLULAR LOCATION: SECRETED. CELL-WALL BOUND.
 CC -1- SIMILARITY: BELONGS TO FAMILY 68 OF GLYCOSYL HYDROLASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: L08445; AAA71925.1;
 DR InterPro: IPR003469; Glyco_hydro.68.
 DR Pfam: PF02435; Glyco_hydro.68; 1.
 KM transferase; Glycosyltransferase; Signal; Cell wall.
 FT SIGNAL 1 969
 FT CHAIN ? LEVANSUCRASE.
 SQ SEQUENCE 969 AA; 103983 MW; D3895B32ACE7735A CRC64;

Query Match 3.9%; Score 91.5; DB 1; Length 969;

Best Local Similarity 19.9%; Pred. No. 16; Mismatches 161; Indels 127; Gaps 20;

Matches 85; Conservative 54; Mismatches 161; Indels 127; Gaps 20;
 QY 53 VQSATOSATSDTPANPLDEHEPELYTALENKTMILNCSALNDIMLA-----CYDPLVNG 108
 DB 155 VEAFTSTASSEAA---DTH-TEVDLKVSNANAMLSKNGIKRISIVEENMTSDIV-A 209
 QY 109 ETPAVIKTRSRJRLDEITQITKGR-POVVOETTPIFLNGNEKMLTKKDAKOLEYAA 167
 DB 210 LTFEELIKALNKWFSD---DAIKGTSTSLTKRLKDIV-----ASFLLKDSKLAVPF 259
 QY 168 KQFTPLSL-SFD-----LDRNTPLMSSRPHPM-----YLPFIHMGKPNRSP 210
 DB 260 KADTIINMPAFNVAQNTKKEIDVWDSMPVQDAKSGVSNMNGVQLVISMAGAPKNKS 319
 QY 211 MPSEHARQFTPNERRAPLTKFOVSKVKAADLMGCTSDLMFGYTGQSHWQIFNKKNSR 270

```

Db 320 NH1LLYRKYGDNDF-----THM-----KNAG 341
Qy 271 PEFVHYOPEIFLTOPYSDLPMDGKYRMIGCAV--HHSNGESAKLSRSMNR---AYLM 325
Db 342 P-----IFGYNMLEDDOQWGSATVNSDGIQLYTKNDTSGGKLMWOLASATLN 392
Qy 326 AGMEMKNTVMRPRIMGRIRFEGSG-----SOP-----DDNPDILIDYGYGD--- 366
Db 393 LAVEDDEVYKSYENDHILFEGDNYHYQSTPKMSYFPDDHNDHNDPDRDNCCLRPHI 452
Qy 367 -----VRELF-----YOLENK-----SNISGYVNRSGKALQLDYVYPLGKIS 407
Db 453 IEDNGSRVILFESNTGDENQGEKQIYKMSNYGDDAFNLKSPNLNKNHNLNLASMAN 512
Qy 408 GYFOIFQ 414
Db 513 GSIGILK 519

```

```

RESULT 13
CATA_DICDI STANDARD; PRT; 496 AA.
ID CATA_DICDI
AC 077229;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CATALASE (EC 1.11.1.6).
GN CATA OR CAT.
OS Dictyostellum discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RA Foote C., Alexander H., Alexander S.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC - FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
PEROXIDE.
CC - CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC - COFACTOR: HEME GROUP (BY SIMILARITY).
CC - SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
CC - SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AF090443; AAC36743.1;
DR HSSP: P00432; 7CAT.
DR DictyDb: DP027272; cata.
DR InterPro: IPR002226; Catalase.
DR Pfam: PF00199; catalase.1.
DR PRINTS: PR00067; CATALASE.
DR ProDom: PD000510; Catalase.1.
DR PROSITE: PS00437; CATALASE.1; 1.
DR PROSITE: PS00438; CATALASE.2; 1.
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Peroxisome.
FT ACT_SITE 54 54 BY SIMILARITY.
FT ACT_SITE 128 128 BY SIMILARITY.
FT BINDING 338 338 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT SITE 494 496 MICROBODY TARGETING SIGNAL (POTENTIAL).
SQ SEQUENCE 496 AA; 55683 MW; 68331B8BFAFD2E6 CRC64;

```

Query Match 3.9%; Score 91; DB 1; Length 496;
Best Local Similarity 23.5%; Pred. No. 7;

```

Matches 56; Conservative 32; Mismatches 94; Indels 56; Gaps 12;
Qy 61 STDANPLDEHEPELTYTALENKTMILNCALNDIMRLACYDPLVNGEPYAVIKTKRSI 120
Db 7 TTSQSPIDNN-----LNSMTAGVNGPILLIDFTLI---DKLAHFDR----- 46
Qy 121 RLDEITWQTIKGRQVYVOETTD-----PIF-----LMGNEKML-T 156
Db 47 RPERVYHA-KGAGAHAFYEYTSDDYKCKAKFLNKVGRKTPFTFESYVGEGKSSDS 105
Qy 157 KDAKQLEYAKQFTPLSLFDDLRNTPLMSSRPHPMYVLPIFMHKKRSPSPESHE 216
Db 106 ERDPR--GFAVKFYTEEG-NPDWVGNTPEVFTRIDPSK---FPDFITQ-KRNPQTNCKD 158
Qy 217 AROFTPNEFRAPELKFOYSYKKAEDLKGTDSDLMFGTYQOSHWOIFNGKNSRPFVY 274
Db 159 PNMEWDFLGOTPESTHQSILFSDR---STPKSYRHHGFSHTLKEVYNAQGRPYWY 212

```

```

RESULT 14
A2MG_HUMAN STANDARD; PRT; 1474 AA.
ID A2MG_HUMAN
AC P01023;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ALPHA-2-MACROGLOBULIN PRECURSOR (ALPHA-2-M).
GN A2M.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85190481; PubMed=2581245;
RA Kan C.-C., Solomon E., Belt K.T., Chain A.C., Hlorens L.R., Fey G.H.;
RT "Nucleotide sequence of cDNA encoding human alpha 2-macroglobulin and
RT assignment of the chromosomal locus.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2282-2286(1985).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT HIS-704.
RX MEDLINE=92246939; PubMed=1374237;
RA Mathijs G., Devriendt K., Cassiman J.-J., van den Berghe H.,
RA Marynen P.;
RL "Structure of the human alpha-2 macroglobulin gene and its promotor.";
RN [3]
RP SEQUENCE OF 24-1474.
RX MEDLINE=84239807; PubMed=6203908;
RA Sottup-Jensen L., Stepanik T.M., Kristensen T., Wierzbicki D.M.,
RA Jones C.M., Loenblad P.B., Magnusson S., Petersen T.E.;
RT "Primary structure of human alpha 2-macroglobulin. V. The complete
RT structure.";
RL J. Biol. Chem. 259:8318-8327(1984).
RN [4]
RP ERRATUM.
RA Sottup-Jensen L., Stepanik T.M., Kristensen T., Wierzbicki D.M.,
RA Jones C.M., Loenblad P.B., Magnusson S., Petersen T.E.;
RL J. Biol. Chem. 260:6500-6500(1985).
RN [5]
RP INHIBITORY SITE.
RX MEDLINE=84030513; PubMed=6195065;
RA Virca G.D., Salvesen G.S., Travis J.;
RT "Human neutrophil elastase and cathepsin G cleavage sites in the bait
RT bait region.";
RL Hoppe-Seyler's Z. Physiol. Chem. 364:1297-1302(1983).
RN [6]
RP INHIBITORY SITE.
RX MEDLINE=81212827; PubMed=6165619;
RA Sottup-Jensen L., Loenblad P.B., Stepanik T.M., Petersen T.E.,
RA Magnusson S., Joernvall H.;
RT "Primary structure of the 'bait' region for proteinases in alpha 2-

```


DB 217 H-----PFTVEEVLKRFEOVTVPKIITILEEMNVSVCGLTYTGKPYGHVTVSICKKYS 273
OY 240 AADLNGTSDLMFGYTGQSHMOIFNGK-NSRPFVHDIQPELFLQPVYSIDLPMQDKVR 298
DB 274 DASDCGEEDS-----DAFCERFSGQLNSHGCIFYQVTKVFOLEKREYEMKLIHTEAQ 325
OY 299 MCGMGAHVHNSNG-ESAKLSNSNR-----AYLMAGMEKNLTVMPRIWGR1-FKEGSG 349
DB 326 IDEEGTVVELTGKQSSSEITRTITKLSFVKVDSHFROG1P-----FGQVLYDQKG 376
OY 350 SOPDDNDPDL-----DYYG-----YGDVRFLYOLENKSNIQGT-----VRYNPRS 389
DB 377 -VPIPKVIFIRGENEANYSNATDEHLQVQFSI-----NTNVMGTSLTVAVNKKDRS 429

RESULT 15
INVO_MOUSE STANDARD; PRT; 467 AA.

ID INVO_MOUSE STANDARD; PRT; 467 AA.
AC P48997;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE INVOLUCRIN.
GN IVL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIH SWISS;
RX MEDLINE=94104476; PubMed=8277848;
RA DJan P., Phillips M., Easley K., Huang E., Simon M., Rice R.H.,
RA Green H.;
RT "The involucrin genes of the mouse and the rat: study of their shared
RT repeats.";
RL Mol. Biol. Evol. 10:1136-1149(1993).
CC -I- FUNCTION: INVOLUCRIN IS A KERATINOCYTE PROTEIN THAT FIRST APPEARS
CC IN THE CELL CYTOSOL, BUT ULTIMATELY BECOMES CROSS-LINKED TO
CC MEMBRANE PROTEINS BY TRANSGLUTAMINASE. ALL THAT RESULTS IN THE
CC FORMATION OF AN INSOLUBLE ENVELOPE BENEATH THE PLASMA MEMBRANE.
CC -I- TISSUE SPECIFICITY: PRESENT IN KERATINOCYTES OF EPIDERMIS AND
CC OTHER STRATIFIED SQUAMOUS EPITHELIA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L28819; AAA39330.1; -
DR MGD: MGI:96626; IYL.
DR InterPro: IPR002360; INVOLUCRIN.
DR PROSITE: PS00795; INVOLUCRIN; 1.
KW Keratinocyte; Repeat.
SQ SEQUENCE 467 AA; 54919 MW; 603E1E51B435737D CRC64;

Query Match 3 8%; Score 90.5; DB 1; Length 467;
Best Local Similarity 19.8%; Pred. No. 7.1; 128; Indels 79; Gaps 14;
Matches 62; Conservative 44; Mismatches 128; Indels 79; Gaps 14;

OY 134 PQVYV-----QETTDPIFLMG-----NEKGLTKKDAQOLEYAAKOFPLSLSFDRN 182
DB 175 PQELHNRQHQEKQDPELHNGOQKPREBKLIPEGKQDELHNGRQEPQDELHNGOK 234
OY 183 NTPLMSSRPANPYVLPIFNHGKPNRSPNTPSHEARQFTPNFEFAPELKFOVSVKAAE 242
DB 235 Q-----KQKLEPELQKQOQKHQPS--EPLEPLKQOQESPE---PELPLGKQOQESPE 285
OY 243 DLMGTSDLMFGYTGQSHMOIFNGKNSRPFVHDIQPELFLQPVYSIDLPMQDKVRIGM 302

DB 286 -----PELQLGKQOQSHEDMAGDQEKQKLIH-----KPELILRKQOQYESP--DEP---LSL 333
OY 303 GAVHNSNGESAKLSRSWNRAYLMAGMEKNLTVMPRIWGR1FKEGSGSOPDDNDPDLIDY 362
DB 334 GKQOHOECQPELQ-----LEEKQHKRPEPELHNGKQOQSHEDMAED----- 377
OY 363 GYGDVRFLYOLENKSNIQGTVRNPRSGKALQLDIYYPILGKISGYFOIFQGYGSL-- 420
DB 378 -----LEEKQKLG-----EPLELQKQOQOQIEREG-----YQGPKSIGOSLQ 416
OY 421 -----IDYNH 425
DB 417 EKASREQQLDYSH 429

Search completed: November 30, 2001, 14:27:03
Job time: 550 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:26:25 ; Search time 77.9 Seconds

(Without alignments)
829,941 Million cell updates/sec

Title: US-09-787-083-4
Perfect score: 2360
Sequence: 1 MKVSLSTLTSLTILPCFALTA.....YNHRTSRGVLMLNDMMGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_17:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mhc:*
9: sp.organelle:*
10: sp.phage:*
11: sp.plant:*
12: sp-rodent:*
13: sp-virus:*
14: sp-vertebrate:*
15: sp-unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	815	34.5	382	2	09K0U7
2	815	34.5	409	2	09JY21
3	373.5	15.8	306	2	09C122
4	366	15.5	329	2	09PMU8
5	360.5	15.3	292	2	09S1D7
6	360	15.3	289	2	09L6N9
7	356.5	15.1	292	2	09Z4N8
8	342.5	14.5	297	2	032349
9	246.5	10.4	355	2	025241
10	240.5	10.2	355	2	09ZLX5
11	121	5.1	278	2	09YB53
12	109.5	4.6	824	2	09F216
13	109	4.6	821	2	059241
14	106.5	4.5	602	11	063485
15	106	4.5	798	4	09UN32
16	102.5	4.5	798	4	09UBK2
17	102.5	4.3	1686	4	000443
18	102.5	4.3	1686	4	000443
19	99.5	4.2	1509	11	061194

20	98.5	4.2	467	4	09UM5	09um5 homo sapien
21	98.5	4.2	467	4	09BRY2	09bry2 homo sapien
22	98.5	4.2	576	2	09L115	09l115 streptomyc
23	98.5	4.2	788	10	09SCV5	09scv5 arabidopsis
24	98.5	4.2	1658	11	061182	061182 mus musculu
25	98	4.2	435	2	09X9C0	09x9c0 streptococc
26	98	4.2	871	4	09Y5C3	09y5c3 homo sapien
27	98	4.2	938	4	09Y5F7	09y5f7 homo sapien
28	97.5	4.1	797	2	09RDM6	09rdm6 lactobacill
29	97	4.1	761	12	09WT88	09wt88 lt virus. o
30	96	4.1	422	2	09RC26	09rc26 streptomyc
31	95.5	4.0	389	2	074515	074515 synechocyst
32	95	4.0	523	4	09H856	09h856 homo sapien
33	95	4.0	739	5	09Y9E6	09y9e6 drosophila
34	94.5	4.0	476	5	09B160	09b160 caenorhabd1
35	94.5	4.0	530	5	045879	045879 caenorhabd1
36	94.5	4.0	683	2	09A6U7	09a6u7 caulobacter
37	94.5	4.0	749	2	059154	059154 anaerocellu
38	94.5	4.0	772	2	0923V2	0923v2 pseudomonas
39	94.5	4.0	1046	2	084941	084941 streptococc
40	94.5	4.0	1641	2	09PDX7	09pdx7 xylella fas
41	94	4.0	454	11	09E054	09e054 mus musculu
42	94	4.0	455	4	012875	012875 homo sapien
43	94	4.0	765	2	054183	054183 streptomyc
44	93.5	4.0	564	3	094727	094727 schizosacch
45	93.5	4.0	664	3	09C122	09c122 pitomyces s

ALIGNMENTS

RESULT	ID	Query Match	Score	DB 2	Length	382	Query Match	Score	DB 2	Length	382
09K0U7	09K0U7	PRELIMINARY;	PRT;	382	AA.		09K0U7	PRELIMINARY;	PRT;	382	AA.
AC	09K0U7	01-OCT-2000 (TREMBLrel. 15, Created)					AC	09K0U7	01-OCT-2000 (TREMBLrel. 15, Created)		
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)						DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)						DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	PHOSPHOLIPASE A1, PUTATIVE.						DE	PHOSPHOLIPASE A1, PUTATIVE.			
GN	NMB0464.						GN	NMB0464.			
OS	Neisseria meningitidis (serogroup B).						OS	Neisseria meningitidis (serogroup B).			
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.						OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.			
OX	NCBI_TaxID:491.						OX	NCBI_TaxID:491.			
RP	SEQUENCE FROM N.A.						RP	SEQUENCE FROM N.A.			
RC	STRAIN-MC58 / SEROGROUP B;						RC	STRAIN-MC58 / SEROGROUP B;			
RC	MEDLINE-20175755; PubMed=10710307;						RC	MEDLINE-20175755; PubMed=10710307;			
RA	Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J., Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A., Mason T., Clecko A., Parksey D.S., Blair E., Clifton H., Clark E.B., Cotton M.D., Ueberback T.R., Khouri H., Qin H., Yamanehan J., Gili J., Scarlato V., Maignani V., Pizza M., Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappelli R., Venter J.C.; "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58."						RA	Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J., Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A., Mason T., Clecko A., Parksey D.S., Blair E., Clifton H., Clark E.B., Cotton M.D., Ueberback T.R., Khouri H., Qin H., Yamanehan J., Gili J., Scarlato V., Maignani V., Pizza M., Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappelli R., Venter J.C.; "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58."			
RT	Science 287:1809-1815(2000).						RT	Science 287:1809-1815(2000).			
RL	EMBL: AE002403; AAF40901.1; -.						RL	EMBL: AE002403; AAF40901.1; -.			
DR	TIQR: NMB0464.						DR	TIQR: NMB0464.			
DR	InterPro: IPR003187; PLAI.						DR	InterPro: IPR003187; PLAI.			
DR	Pfam: PF02253; PLAI; 1.						DR	Pfam: PF02253; PLAI; 1.			
KW	Complete proteome.						KW	Complete proteome.			
SO	SEQUENCE 382 AA; 42714 MW; B46BA802F062E836 CRC64;						SO	SEQUENCE 382 AA; 42714 MW; B46BA802F062E836 CRC64;			

Query Match	34.5%	Score 815;	DB 2:	Length 382;
Best Local Similarity	44.7%	Pred. No. 6.5e-63;		
Matches 163;	Conservative 67;	Mismatches 111;	Indels 24;	Gaps 8;

OY	141	TTDITFLMNGENKGLT/KKKDAKOLEYAKOCPELSTSDLDLRNN-TPLMSSRPNNPWTLD	199
Db	93	GGDAL-----PADSAGETDIYTPPLSLMTDLDKNDRLGELGYREINNPATLAP	139
OY	200	IFPHGKPNRSFNTPSHEAR-OFTPNFEAPDLKFOYSVKVKAADLMGSDLMFGYTQO	258
Db	140	LMYNNSPYAGSGPTRGTTVOEKGGOOKRAETKLOVFSKSIABEDTEKTRADLMFGYTQO	199
OY	259	SHMOIFN-GKKRSRPFRRHNDYORPEITFLPOPYUSDLPWMCGRKRMIGMGVHNHNSGESAKLSR	317
Db	200	SDMOIYNGRKSAPFRMTDYKPELEFLLOPYKADLPFGGRGLRMIGAGVHNSQNSQSRPESR	259
OY	318	SMNRATYLMAGHEKNLTYMPRIWIKRITKEGSGSOPDDNPILDIYEGIDVRYLTOLENKS	377
Db	260	SMNRITAYAGGMEWCKLTYIPRWVRAADQ-SGDK-NDNPDIAIYMGGDVCKLQYRLANDRO	317
OY	378	NISGTVRNPRPSGALOLDVYVPLGKIGISGYFOIFQGYQCSLDLIDYHNKATSFVGLMLN	457
Db	318	NVYSVLRNPNRPTGCALEAATPTPIKCKLKGUVNMGFGVGBESLDLYHNKNGIGIGIMFN	377
OY	438	DMKGL 442	
Db	378	DLGDI 382	

RESULT	2			
09JUT1				
ID	09JUT1	PRELIMINARY;	PRT;	409 AA.
AC	09JUT1;			
DT	01-OCT-2000 (TREMBlrel. 15, Created)			
DT	01-OCT-2000 (TREMBlrel. 15, Last sequence update)			
DT	01-JUN-2001 (TREMBlrel. 17, Last annotation update)			
DE	PUTATIVE PHOPHOLIPASE.			
GN	NMA2021.			
OS	Neisseria meningitidis (serogroup A).			
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.			
OX	NCBI_TaxID=65699;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=22491 / SEROGROUP A / SEROTYPE 4A;			
RX	MEDLINE=20222556; PubMed=10761919;			
RA	Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,			
RA	Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,			
RA	Davies R.M., Davis P., Devlin K., Feltham T., Hamlin N., Holtroyd S.,			
RA	Jagels K., Leather S., Moule S., Mungall K., Quill M.A.,			
RA	Randread M.A., Rutherford K.M., Simmonds M., Skelton J.,			
RA	Whitehead S., Spratt B.G., Barrall B.G.;			
RT	"Complete DNA sequence of a serogroup A strain of Neisseria			
RT	meningitidis 22491."			
RL	Nature 404:502-506(2000).			
DR	EMBL; AL162757; CAB85240.1; -.			
DR	InterPro; IPR003187; PLAI.			
DR	Pfam; PF00253; PLAI; 1.			
KW	Complete proteome.			
SO	SEQUENCE 409 AA; 45862 MW; CD6585B064D01A41 CRC64;			

[illegible]

QY	259	SHMOEEN-GKNSRPRRYNDXOPELEFETLORVYSDLPBQKGYMIGMGAVHNHNSGSAKLRS	31.7
Db	227	SDMOQYINQGRKSABPRNTDYKPELEFETLORVYKADLPBEGRLMLAGFVHOSNGOSRPSER	28.6
QY	318	SMNRAYVLAAGAEWKNLTVMPRIWGRIFKEGSGSOPDDNPDLIDYVYGQDVRFLVQLENKS	37.7
Db	287	SMNRVYLAAGAEWCKLTVIAPVWVRAFQD-SCGR-NDNPDIADYMGVGDVKYQRLANDRQ	34.4
QY	378	NISGTAVRYNPPSGGALDLDVYVPLGKGISCTYFDIPOGIGOSLIDYVHHEATSRFVGMLN	43.7
Db	345	NVYSWLRNPKPTGGALEAAATYPIKGLKGVNKGFGHGESLIDYVHNKHNGICIGIMFN	40.4
QY	438	DMWGL	44.2
Db	405	DLDDGI	40.9

```

RESULT      3
09CL22      PRELIMINARY;      PRT;      306 AA.
AC      09CL22:
DT      01-JUN-2001 (TREMBLrel. 17, Created)
DT      01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT      01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE      HYPOTHETICAL PROTEIN PM1426.
GN      PM1426.
OS      Pasteurella multocida.
OC      Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC      Pasteurella
OX      NCBI_TaxID=747;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=PM70;
RX      MEDLINE=21145866; Pubmed=11248100;
RA      May B.-J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT      "Complete genomic sequence of Pasteurella multocida PM70."
RL      Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR      EMBL; AE006179; AAK03510.1; -.
DR      InterPro: IPR003187; PLA1.
DR      Pfam; PF02253; PLA1; 1.
KW      Hypothetical protein; Complete proteome.
SQ      SEQUENCE      306 AA;      35580 MW;      EAF3DE8C1C22B26E CRC64;

```

	Query Match	15.86;	Score 373.5;	DB 2;	Length 306;
	Best Local Similarity	37.66;	Pred. No. 1.Gc-24;		
	Matches 86;	Conservative 43;	Matches 91;	Indels 9;	Gaps
Oy	214 SHEAQTPEPNE-FRARELKFOVSVKVAKAEDGDSIDMFGYQOOSHWOLFNGKNRRPF	272			
Dd	81 TYSNHFLEFKRKRTQODELTKFIKISALPLMRGIILNNVSYLASAYQKSMFOLDNSDDSPF	140			
Oy	273 RVHDOPEIFLTROPYSDLPMDKGVRMGAVHNHSNC--ESAKTSRWNPAYLLMAGMEM	330			
Dd	141 RETNTEPOFLAMKTQYSLPLPGMTLDQVEGTINHGSGNRDRAEKLSSWNRPLYRASAIK	200			
Oy	331 KNLTVPRIMGRIRFEKSSGSODDPNDPIILDYGYGDYRF-PLOELNKSNISGTGVRYNPRS	388			
Dd	201 QNMVLDEIKRPWRIRPK---AKNDNDPDITKRGHFDVALGYVYHDHQFKLSG--HYNPIS	255			
Oy	330 GKALGLDYDYPLKGISGYFOIFOGYGQSILIDYNHEATSEFGVGLMND	438			
Dd	256 NKGGLEASYSPITFKNIREFYQYNGYGESLIDYQORIQRIQIGISLNN	304			
RESULT	4				
Q9PMU8					
Q9PMU8	PRELIMINARY;	PRT;	329 AA.		
AC	Q9PMU8:				
DT	01-OCT-2000 (TREMBLrel. 15,	Created)			
DT	01-OCT-2000 (TREMBLrel. 15,	Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17,	Last annotation update)			

DE PHOSPHOLIPASE A (EC 3.1.1.32).
CN PLDA OR CUI351.
OS Campylobacter jejuni.
OC Bacteria: Proteobacteria; epsilon subdivision; Campylobacter group;
OX Campylobacter.
NCBI_Taxid=197;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=NTC 11158;
RA MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Kelley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Felwell T., Holtroyd S.,
RA Jørgensen K., Kariyasekera A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rabsdram M.A., Rutherford K.M., Van Vleet A.H.M.,
RA Whitehead S., Barrell B.G.,
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences."
RL Nature 403:665-668(2000).
DR EMBL: AL139078; CAB3778.1; -.
DR InterPro: IPR003187; PLAI.
DR Pfam: PF02253; PLAI.1
DR PRINTS: PRO1486; PHPLIPASEA1.
KW Complete proteome.
SQ SEQUENCE 329 AA; 38880 MW; A32731F2B751AC4 CRC64;

Query Match 15.5%; Score 366; DB 2; Length 329;
Best Local Similarity 33.2%; Pred. No. 8.2e-24;
Matches 105; Conservative 48; Mismatches 125; Indels 38; Gaps 13;

OY 130 IKGAPVYVYVETTPPIFLMGNEKGLTKKDAKOLEYAKOFTPLSLSDLDNRNNT--PLW 187
DB 43 LKNSSVLSQEQNNSSQATOTQNTITIKKEKQ-----DPSRLALAVYLGENSEFNLG 96
OY 188 SSRPNPVPYVPIFMHGKPNRSPTPSHEARQTPNEFRABELKROVSVKAAEDLWGT 247
DB 97 IS-SYKMYFLP-FAYSPNSLGVNNKSEA-----KQLSVKRLKLEFNLGL 141
OY 248 DSDLMFGTQOOSHQIENGKNSRPFRVHDYQPEIFLTPYV-SDLPMGKYMIGMAVH 306
DB 142 DEKYIATQTSWQIY--EHSSPRETNQPEEFIDLPFLKDYEFNNLR--VGILH 196
OY 307 HSNSESAR--LSRSMNRAYLMAGMEKMLTYVPRIMGIFEGSGSDPDNDLIDYGY 364
DB 197 ESNKGDENLQSRMNRITVSTALYKFLVPRIMYRI---PEKKKDDNPAILHYGN 253
OY 365 GDVAFELYOLEKNSISGTVRYNPR--SGKALQLDYVPL-GKISGYFOIFQGYGSLI 421
DB 254 EDVNLAY-LGDYFINLMLRNKLFHNKGAIQYDLYGDIENNIGIYWLQYFNGYGESLI 312
OY 422 DYNHEATSFVGLMLN 437
DB 313 DYNKHLQRLSTGFLIS 328

RESULT 5
OQSID7 PRELIMINARY; PRT; 292 AA.
ID OQSID7
AC OQSID7;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE PHOSPHOLIPASE A.
GN PLDA.
OS Yersinia pseudotuberculosis.
OC Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX Yersinia.
NCBI_Taxid=633;
RN SEQUENCE FROM N.A.
RC STRAIN=VPIII P1B1;
RA Kariyasekera A.V., Winzeler E.A., Williams K.J., Oyston P.C.,
RA Tibball R.W., Wren B.W.;

RT "Biochip-based Signature-Tagged Mutagenesis: identification and
RT characterisation of Y. pseudotuberculosis gene plda essential for
RT virulence in mice."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ245393; CAB51586.1; -.
DR InterPro: IPR003187; PLAI.
DR Pfam: PF02253; PLAI.1
DR PRINTS: PRO1486; PHPLIPASEA1.
SQ SEQUENCE 292 AA; 33758 MW; 8E712D908ACB6BA5 CRC64;

Query Match 15.3%; Score 360.5; DB 2; Length 292;
Best Local Similarity 39.2%; Pred. No. 2.1e-23;
Matches 83; Conservative 34; Mismatches 86; Indels 9; Gaps 5;

OY 229 EIKFQSVKVKAAEDLWGTSDLMFGYTOOSHQIENGKNSRPFRVHDYQPEIFLTPYV- 287
DB 86 EVKFLSLAFPIFMRGICAGNSLIGASYTQSRWQASNSESSPFRETNYEPOLFLAWSTD 145
OY 288 YSDLPWDGKVRMIGMAVHNSGESAKLSRSNRRAYLMAGMEKMLTYVPRIMGRIFREG 347
DB 146 YELAGW--TFREVERGFPHQSGKADPTSRSMNRAYTRYMAQGNLEIDLKFWYRIPESD 203
OY 348 SGSDPDNDLIDYGYGVDFRFLYOL-ENKSNISGTVRYNPRSGKALQLDYVPLGKI 406
DB 204 S---KDDNPDIKKYMGYRLKAGYALGDSVFSLDG--RNNMTGICGAEKMGSTYITKHV 258
OY 407 SGYFOIFQGYGSLIDYNHEATSFVGLMLN 438
DB 259 RYTGVSFGSGEMIDYNTROTGVGIMLND 290

RESULT 6
O9L6N9 PRELIMINARY; PRT; 289 AA.
ID O9L6N9
AC O9L6N9;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE PLDA PROTEIN.
GN Salmonella typhimurium LT2.
OC Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX Salmonella.
NCBI_Taxid=99287;
RN SEQUENCE FROM N.A.
RC STRAIN=SGS1412;
RA WashU;
RT "The Salmonella typhimurium Genome Sequencing Project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SGS1412;
RA Waterston R.
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF233324; AAF33435.1; -.
DR InterPro: IPR003187; PLAI.
DR Pfam: PF02253; PLAI.1
SQ SEQUENCE 289 AA; 32967 MW; DA97F5E1651C49C6 CRC64;

Query Match 15.3%; Score 360; DB 2; Length 289;
Best Local Similarity 39.7%; Pred. No. 2.3e-23;
Matches 85; Conservative 29; Mismatches 90; Indels 10; Gaps 4;

OY 226 RABELKFOYSVVKAAEDLWGTSDLMFGYTOOSHQIENGKNSRPFRVHDYQPEIFLTPQ 285
DB 83 RKDEVKFOLSLAFPIFMRGICAGNSLIGASYTQSRWQASNSESSPFRETNYEPOLFLGCF 142
OY 286 PV-YSDLPWDGKVRMIGMAVHNSGESAKLSRSNRRAYLMAGMEKMLTYVPRIMGRIF 344
DB 143 ATDYRFAGW--TLRDVEKGYNHDSNGRSDPTSRSMNRILTRYLMAGNMGWIVYKRWYI- 199

```
OY 345 KEGSGSQPDNDPILDYGGVRELYQLEENKSNISGTVNPRSGKALQLDVYVPLGK 404
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 200 -----GSPDNDPITFKYMGYQKIGYHL-GEAVLSAKGQYNNMTGGAEGVLSYVTK 253
OY 405 GISGYFOIRGCGSLIDYHNEATSFVGLMND 438
      : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 254 HVRLTYGVYSGESLIDYFNQTRVGVGLMND 287

RESULT 7
O924N8 PRELIMINARY; PRT: 292 AA.
AC O924N8;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
DE OUTER MEMBRANE PHOSPHOLIPASE A PRECURSOR (EC 3.1.1.32).
GN PLDA.
OS Enterobacter agglomerans.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Pantoea.
OX NCBI_TaxID=549;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94131966; PubMed=8300539;
RA Brok R.G., Brinkman E., van Boxtel R., Bekkers A.C., Verheij H.M.,
RA Tomassen J.;
RT "Molecular characterization of enterobacterial plda genes encoding
RT outer membrane phospholipase A. ";
RN J. Bacteriol. 176:861-870(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120145; PubMed=9921577;
RA Brok R.G., Boots A.P., Dekker N., Verheij H.M., Tomassen J.;
RT "Sequence comparison of outer membrane phospholipases A: implications
RT for structure and for the catalytic mechanism.";
RL Res. Microbiol. 149:703-710(1998).
DR EMBL: AF034414; AAD03498.1;
DR InterPro: IPR003187; PLA1.
DR Pfam: PF02253; PLA1; 1.
DR PRINTS: PR01486; PHPLIPASEA1.
KW Signal; Hydrolase.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 292 OUTER MEMBRANE PHOSPHOLIPASE A.
SQ SEQUENCE 292 AA; 33719 MW; B75516D093B2BEEA CRC64;

Query Match 15.1%; Score 356.5; DB 2; Length 292;
Best Local Similarity 40.1%; Pred. No. 4.7e-23;
Matches 85; Conservative 31; Mismatches 87; Indels 9; Gaps 5;

OY 229 ELKFOYSVKKAAEDLWGTSDFGYSQSHQIFNGKNSRPFVADYQPEIFLQPV- 267
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 86 EVFQSLGPIRGIAIGDNLGASTQSRMWAQNSDSSPFRETNPPQIFLAMATD 145
OY 288 YSDLPWDGKVRMIGAVHNSGSAKLSSRNKRAYLMAGMEKNLTVMPRIKRFEGK 347
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 146 YELAGW--TFREVEFGYHQSNGKADPTSRSDRVYTRLMAQGNLEIDLPKPYRIEED 203
OY 348 SGSGQPDNDPILDYGGVRELYQLE-ENKSNISGTVNPRSGKALQLDVYVPLGKI 406
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 204 S---KQDNPINKYGRYRLKVGVALGESVFSIDG--RYNMNMGYGAEGMGSYPIRKV 258
OY 407 SGYFOIRGCGSLIDYHNEATSFVGLMND 438
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 259 RFTYVFSGIGESMIDYFNQTRVGVGLMND 290

RESULT 8
O32349 PRELIMINARY; PRT: 297 AA.
AC O32349;
```

```
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
DE PHOSPHOLIPASE A.
GN PLDA.
OS Campylobacter coli.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=195;
RN [1]
RP SEQUENCE FROM N.A.
RX STRATIN-04585;
RX MEDLINE=97230284; PubMed=9119448;
RA Grant K.A., Belandia I., Dekker N., Richardson P.T., Park S.F.;
RT "Molecular characterization of plda, the structural gene for a
RT phospholipase A from Campylobacter coli, and its contribution to cell-
RT associated hemolysis ";
RL Infect. Immun. 65:1172-1180(1997).
DR EMBL: Y11031; CA471915.1;
DR InterPro: IPR003187; PLA1.
DR Pfam: PF02253; PLA1; 1.
DR PRINTS: PR01486; PHPLIPASEA1.
SQ SEQUENCE 297 AA; 34998 MW; 04B54A7BCA3764CE CRC64;

Query Match 14.5%; Score 342.5; DB 2; Length 297;
Best Local Similarity 32.4%; Pred. No. 8e-22;
Matches 97; Conservative 37; Mismatches 108; Indels 57; Gaps 11;

OY 178 DLDNRNTPLMSSRPBHPNMYVLPIFMGKPNRSPT-----P 213
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 16 DLKENNASLIRIKHEQ-----NTKPTSTKEDFSRIALANYGENSSFNPLGIS 66
OY 214 SHEARQTP-----NEFRAPELKFOYSVKKAAEDLWGTSDFGYSQSHQIF 264
      | : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 67 SYKNTYLPFRAYSFGSLGKRNKTEKFKLSIKRLEFDLGLGEKTYVGTOTSMQ-- 124
OY 265 NGRNSRPFVADYQPEIFLQPV-YSDLPWDGKVRMIGAVHNSGSAK--LSRSWR 321
      | : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 125 NYKHSSPFRETNQPEFVQDLPCHFEDYKFLNLR--VGILHSNGKGDENLESRSWR 181
OY 322 AYLMAGMEKNLTVMPRIKRFEGSGSQPDNDPILDYGGVRELYQLEENKSNISG 361
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 182 IVASSVFLYQRFLEVPRIWRI--PENSDDDNPELTHMGNEFDIN-IGSLGNDYFINL 237
OY 382 TVRYNP--RSQKALQLDVYVPL-GRGISGYFOIRGCGSLIDYHNEATSFVGLMNL 437
      | : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 238 MLRNNDLFDHDKGAVQVIGDIFDNGIYIYLOTFNGYGSGLIDYKRLQRLSTAFLLIS 296

RESULT 9
O25241 PRELIMINARY; PRT: 355 AA.
AC O25241;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
DE PHOSPHOLIPASE A1 PRECURSOR (DR-PHOSPHOLIPASE A).
GN HPO499.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RX STRATIN-26595 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
```

RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Mallin E.,
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
 RA Venter J.C.:
 RT "The complete genome sequence of the gastric pathogen *Helicobacter*
 RT *pylori*.";
 RL Nature 388:539-547(1997).
 DR EMBL: AE000564; AAD07564.1; -
 DR TIGR: HP0499; -
 DR InterPro: IPR003187; PLA1.
 DR Pfam: PF02253; PLA1; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 355 AA; 42486 MW; 461960F15E65AB0A CRC64;

Query Match 10.4%; Score 246.5; DB 2; Length 355;
 Best Local Similarity 25.1%; Pred. No. 2.5e-13;
 Matches 82; Conservative 41; Mismatches 117; Indels 87; Gaps 11;

QY 157 KKDAKOLEYAKOFTPLSLSFDDLDRNTPLMSSRPHNPMVYLPIFMHGKPRSPNTSHE 216
 DB 69 KKYLMMDYLGTYFLPYHSF-----TPIFGWYHPINP----- 102
 QY 217 AROETPNEFRAPELKFOVSVKKAEDLWGTDSDLMFGYTOOSHMOJFNKNSRPFVHD 276
 DB 103 ---YORNEF-----KFOISFRVPRVFRHILMTKGLTYLATYOTDMFOIYNDSQAPMRMN 154
 QY 277 YQPEIFLTQPYVSDLPMDGK---RMIGCAVHNSG-ESAKLSRSNRAVLAGMEWKN 332
 DB 155 FMPELIYVYPI-NKPPGKIGNESEIWMQHSNGVGAOCYQPFNK---EGNPENQ 209
 QY 333 LTWVPRF-----WGRIFKSGSGSOP-----DNDPD 357
 DB 210 PPGOPVTVKDYNGKDVKMGCRSVSAGQRPVFLWMEKGLKIMVAYWPYVQDSHPN 269
 QY 358 ILDYGYGADVREFY-----OLENKSNIQTVYRNPNSGKALQLDVYVPLGKISGYF 410
 DB 270 LIDWYGKGNKIDYRGRHHELDYDIFQYWRD--RHWGAPRLGTYTINPFVGIYA 327
 QY 411 QIFQGIQOSLIDVNHETSRFGVGLMLN 437
 DB 328 QWFMNGYDGLYEDVDFSNRIGVGIRLN 354

RESULT 10
 Q9ZLX5 PRELIMINARY; PRT; 355 AA.
 AC Q9ZLX5;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE PUTATIVE PHOSPHOLIPASE A1.
 GN PLDA OR JHP0451.
 OS *Helicobacter pylori* J99 (Campylobacter *pylori* J99).
 OC Bacteria; Proteobacteria; epsilon subdivision; *Helicobacter* group;
 OC *Helicobacter*.
 NC NCBL_Taxid=85963;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99120557; Pubmed=9923682;
 RA Alm R.A., Ling L.-S.L., Moll D.T., King B.L., Brown E.D., Dolg P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
 RA Tummlino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RT gastric pathogen *Helicobacter pylori*.";
 RL Nature 397:176-180(1999).
 DR EMBL: AE001479; AAD06029.1; -
 DR InterPro: IPR003187; PLA1.
 DR Pfam: PF02253; PLA1; 1.
 KW Complete proteome.
 SQ SEQUENCE 355 AA; 42329 MW; B3CE9810EBA2FDC9 CRC64;

Query Match 10.2%; Score 240.5; DB 2; Length 355;
 Best Local Similarity 25.1%; Pred. No. 8.3e-13;
 Matches 83; Conservative 44; Mismatches 109; Indels 95; Gaps 13;

QY 157 KKDAKOLEYAKOFTPLSLSFDDLDRNTPLMSSRPHNPMVYLPIFMHGKPRSPNTSHE 216
 DB 69 KKYLMMDYLGTYFLPYHSF-----TPIFGWYHPINP----- 102
 QY 217 AROETPNEFRAPELKFOVSVKKAEDLWGTDSDLMFGYTOOSHMOJFNKNSRPFVHD 276
 DB 103 ---YORNEF-----KFOISFRVPRVFRHILMTKGLTYLATYOTDMFOIYNDSQAPMRMN 154
 QY 277 YQPEIFLTQPYVSDLPMDGK---RMIGCAVHNSG-ESAKLSRSNR----- 321
 DB 155 FMPELIYVYPI-NKPPGKIGNESEIWMQHSNGVGAOCYQPFNKEGNPENQPFQ 213
 QY 322 -----AYLMAG-----MENK--NTWVPRFGRIFKSGSGSOPDD- 354
 DB 214 PVIYKDYNGKDVKMGCRSVSAGNALCFVLWMEKGLKIMVAYWPY-----PYDQ 265
 QY 355 -NPDILYGYGADVREFY-----OLENKSNIQTVYRNPNSGKALQLDVYVPLGKI 406
 DB 266 SNPQIDWYGKGNKIDYRGRHHELDYDIFQYWRD--RHWGAPRLGTYTINPFV 323
 QY 407 SGYFOIQOSLIDVNHETSRFGVGLMLN 437
 DB 324 GIVQWFMNGYDGLYEDVDFSNRIGVGIRLN 354

RESULT 11
 Q9XB53 PRELIMINARY; PRT; 278 AA.
 AC Q9XB53;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE KDOT.
 GN *Erwinia carotovora*.
 OS *Erwinia carotovora*.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Pectobacterium.
 NC NCBL_Taxid=554;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 39048; GS101;
 RX MEDLINE=98065591; Pubmed=9402024;
 RA McGowan S.J., Sebaihia M., O'Leary S., Hardie K.R., Williams P.,
 RA Stewart G.S., Bycroft B.W., Salmond G.P.;
 RT "Analysis of the carbenpenem gene cluster of *Erwinia carotovora*:
 RT definition of the antibiotic biosynthetic genes and evidence for a
 RT novel beta-lactam resistance mechanism.";
 RL Mol. Microbiol. 26:545-556(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 39048; GS101;
 RX MEDLINE=98276484; Pubmed=9614345;
 RA McGowan S.J.; Bycroft B.W., Salmond G.P.;
 RT "Bacterial production of carbenpenems and clavams: evolution of beta-
 RT lactam antibiotic pathways.";
 RL Trends Microbiol. 6:203-208(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 39048; GS101;
 RA McGowan S.J.;
 RL submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U17224; AAD38237.1; -
 SQ SEQUENCE 278 AA; 31341 MW; 63769F4F3550E1B7 CRC64;

Query Match 5.1%; Score 121; DB 2; Length 278;
 Best Local Similarity 22.0%; Pred. No. 0.016;
 Matches 71; Conservative 49; Mismatches 125; Indels 78; Gaps 15;

Oy	70	EHEBELYTTALEENKTMMLNCSALNODIMRLACDYLTHGETRAVIKTKSIRLDEITMO	128
Dd	10	EHAATLTTELTKRKFLLEQJFTTPNOYMTYSHIDRIYVG--GIMPVGEITFDDIGIKO	66
Oy	129	-----TIKGROVVYQETTPRIFLMGNCKMILRKDAKOLEYAAKOFETPL	173
Dd	67	FGVAYFLERRELGLINIGPRAKIYIDGS---YEVCNEEALYYGAKKALAFS-----	116
Oy	174	SLSFDLDNRNTPLWSSRPHPMYVLPIFMHGKPRSPFPSSHAROTPPNEFRAPELKEQ	233
Dd	117	--SLDSAKPAKLYYNASAPAHAVPFTRITIQDDAIKAPLGDKVKCNKRICKIYLPEV---	171
Oy	234	VSVYVKAADIEDMGTSDDLMEGYT---QQSHWOJFNKNKSNRPFRVHDYQEPETPLTOPYSD	290
Dd	172	----VEITCQ-----LSMGIPLTAEGSNW-----NSMPRTTHERRMEYV----YFD	209
Oy	291	LPMOGKYRMIGMGAVNHVS-----NGESAKLSRSNNRAYLMAAGENKKLLTYMPRLNGRIEK	345
Dd	210	MAEDTLTFHM-MGPHEHTRHLLVMHNEDAVISPSWS---IHGTGVTKNYAF--IWGM --	260
Oy	346	EGSGOPDDNPDIUDYYGGCVDR	368
Dd	261	-GENLTEDD-----MDHTAMEDLR	278
RESULT	12		
O9F216	ID	O9F216	PRT; 824 AA.
AC	O9F216;	PRELIMINARY;	
DT	01-MAR-2001 (TREMBLrel. 16, Created)		
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	CeLUlASE.		
OS	Bacillus sp.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;		
OC	Bacillus/Staphylococcus group; Bacillus.		
OX	NCB1_Taxid=1409;		
RN	[1]		
SEQUENCE FROM N.A.			
RC	STRAIN-KSM-S237;		
RX	MEDLINE=21036886; Pubmed=11193393;		
RA	Hakemada Y., Hatada Y., Koike K., Yoshimatsu T., Kawai S.,		
RA	Kobayashi T., Ito S.;		
RT	"Reduced amino acid sequence and possible catalytic residues of a		
RT	thermostable, alkaline cellulase from an alkaliophilic Bacillus		
RT	strain".		
RL	Biosci. Biotechnol. Biochem. 64:2281-2289(2000).		
DR	EMBL; AB0184420; BAB19360.1; "-		
DR	InterPro: IPRO01547; Glyco_hydro_F5.		
DR	Pfam; PF00150; cellulase; 1.		
DR	PROSITE; PS00659; GLYCOSTYL-HYDROL-F5; UNKNOWN_1.		
SEQ	SEQUENCE 824 AA; 91564 MW; 65FA940FE1D729B9 CRC64;		

```

Query Match 4 63: Score 109.5: DB 2: Length 824:
Best Local Similarity 19.7%: Pred 0.76:
Matches 91: Conservative 57: Mismatches 146: Indels 167: Gaps 27:

OY 30 PVAIFYEVRSKND-----LGODNELLIGVQATGASDTFANPL-DEHEPELY----TT 78
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 25 PAALAAEENTTRDNEKPHLGDNDNVRRPSEAGALQLOEYDGMOTLVDOHEKIQLRGMSTH 84
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 79 ALENTKMLINCSEAL-----NODIMRLACYDTLVHGE-----TPAVIKTK-----RSI 120
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 85 GLQWPEELINDNAYKALSDMDSMNIRLAMY-----VGEAGYATNELLIKQRIYDGLALAI 140
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 121 RLDEIET---WQI-----IKKRPQVYQETTDI-----TF 146
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 141 ENDMETIVDMVHVAEGDPPROPVYAGAKDFEETIALALYPNNPIIYELANEPSSNNNGAG 200
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 147 LMGNKGMILTKKDAKOLEVYAAKQFTPLSLSPDLDRN-----NTPLWSSRP-----H 192
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Db 201 IPNNEGKAKV-----EYAPIYEMLRKSSNADNDIIIVGSPWMSGRPLDANDPIDH 255
QY 193 NPMYVLPITMGHGKPNRKSPTNPSHEAROTFPMFEAPBLK-----FOVSQVYKKAEDIMCT 247
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 256 HTMYVHHFYTGSSHAASSTYSPSE-----TPNSEGVMNMTRYALEGVAVFATE--WGT 308
QY 248 -----DSDLMEGYTQO-----SHQIQENGKNS-----RPRPV----- 274
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 309 SQASGSGPYDEADVWIEFLNENNISMANNSLTN-KNEVSGATPFELDKSNATINDPG 367
QY 275 --HDYQ--EIFLT-----QPV-----YSDLPW-----DGKVMIGMAVHHSNGE 311
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 368 PDHWAPPELELSGEYVVARIKGVNEYEDRTKTKYKLMLPFNDQTKO--GFGVNSDSPNK 425
QY 312 SAKTSLSRNNRAYLMAAGM-----WKRLTYMPRIWG 342
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 426 ELIAVDNENNTLKYSGLDVNSDVSDDGNFWANARLSANGWGK 466

```

Result	13			
059241				
ID	059241	PRELIMINARY;	PRT;	821 AA.
AC	059241;			
DT	01-NOV-1996 (TReMBLrel. 01, Created)			
DT	01-NOV-1996 (TReMBLrel. 01, Last sequence update)			
DT	01-JUN-2001 (TReMBLrel. 17, Last annotation update)			
DE	ENO-1,4-BETA-GLUCANASE (EC 3.2.1.4) (CELLULASE) (ENDOGLUCANASE)			
DE	(CARBOXYMETHYL CELLULASE).			
OS	Bacillus sp.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;			
OC	Bacillus/Staphylococcus group; Bacillus.			
OX	NCBI_TaxID-1409;			
LN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-KSM-64;			
RA	Suntomo N., Ozaki K., Ito S.;			
RL	Submitted (May-1993) to the EMBL/GenBank/DBJ databases.			
CC	-1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC			
CC	LINKAGES IN CELLULOSE.			
CC	EMBL; M84963; AAA73189.1; -.			
DR	HSSP; O85465; 1A3H.			
DR	InterPro; IPR001547; Glyco_hydro_F5.			
DR	Pfam; PF00150; cellulase; 1			
DR	PROSITE; PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN.1.			
DR	HydroLase; Glycosidase.			
SW	SEQUENCE 821 AA; 90910 MW; 73D438FEF0D40B5C CRC64;			

Query Match	4.6%	Score 109;	DB 2;	Length 821;
Best Local Similarity	19.7%;	Pred. No. 0.83;		
Matches	91; Conservative	51; Mismatches	150; Indels	170; Gaps
QY	30	PVAFAFDEVRSKND-----LGODNELLIGVQSMQASSTDPANL-DEHEPELY-----TT	78	
DB	24	PALAAEGTREDNREKHLIGNDVKRPSAGALQLOEVDGOMTLVDQHEKTLQLRSMSTH	83	
QY	79	ALENTMTLINCAL-----NODIWRILACY-DTLVHGEPYAVIKKRSIRLDETI----	126	
DB	84	GIQWPEELINDAAYAKALANDWESNMIRLAMYGENGYANPELISRSYVTKGIDLIENDM	143	
QY	127	----WQT-----IKGPRVYVQETTD-----IFLMGN	150	
DB	144	YVIVDMHVAHPGDDPRDYVYAGAEDEFRIADIALYPPNPHIILELANEPSSNNNGAGIPNN	203	
QY	151	EKGMLTKRAKQLEFAAKQFTPLSLSFDDRN-----NTPLMSSRP-----HNPMY	196	
DB	204	EEGMAVAVK-----EYADAPIEMLRDSGNDADDNIIVIGSPNWSQRPDLADNPIDHDHTMY	258	
QY	197	VLPDIF--MHGKPNRS--PMTPSHEARQFPNPEFRAPELKFQVSVYKKAADLMGT-----	247	
DB	259	TVHFYTGSHAASTESTPEPTPNSERGNVMSNTRYA-----LENGCAVAFATE--WGTSOANG	312	
QY	248	-----DSDLFEGYTOQ-----SHWQINGRNS-----RPERVH-----	275	

```

Db 313 DCGPYFDEADYIEFLMENNISMANNSLTN-KNEVSAFFPEELGKSNATSLDGPDPYV 371
Oy 276 -----DYQPEIFLTPYSDLPW---DGKVMIGMGAHHNSGES- 312
Db 372 VPEELSLSGEYVARIRKGVNEP---IDRTKYFLVMDPFGNGFGV-----NGDSP 422
Oy 313 -----AKLSRSMNRAYLMAGMEKNLTPMPRIMCR 342
Db 423 VEDVYIENAGALKLGLSDASNDVSEGNMAMARLSADGCK 464

```

RESULT 14

```

ID 063485 PRELIMINARY; PRT: 602 AA.
AC 063485;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE C-RAF ACTIVATED ONCOGENE FUSION PROTEIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=67172791; PubMed=3550433;
RA Ishikawa F., Takaku F., Nagao M., Sugimura T.;
RT "Rat c-raf oncogene activation by a rearrangement that produces a
RT fused protein."
RL Mol. Cell. Biol. 7:1226-1232(1987).
CC - SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: M15428; AAA42002.1; -
DR InterPro: IPR000719; Euk_kinase.
DR InterPro: IPR002290; Ser_thr_kin_actsite.
DR Pfam: PF00069; kinase.1.
DR SMART: SM00221; STYKc.1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
DR PROSITE: PS00101; PROTEIN_KINASE_DOM.1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST.1.
DR APP-Binding: Kinase; Oncogene; Serine/threonine-protein kinase;
KW Transferrase.
SQ SEQUENCE 602 AA; 69209 MW; 88B7BFA90FFB02AC CRC64;

```

Query Match 4.5%; Score 106.5; DB 11; Length 602;
 Best Local Similarity 19.0%; Pred. No. 0.88;
 Matches 92; Conservative 74; Mismatches 172; Indels 145; Gaps 24;

```

Oy 23 QQAQVPPVAVFVDEVRKNDLGONELLIGVQASSTQSTFANPLDHEPELYTTALEN 82
Db 81 ELEFLNNQVYVLEKNEKLETAODRNL--GIQSOFTRAKEE---LEAEKRDILRT--N 131
Oy 83 KTMNLINCSALNODIMRLACYDTLVHGEFPAVYIKTRSRIRDETIWQ--TIKGPQVYVOE 140
Db 132 ERLSQEVEYLETEDVKRL--NEKLESNT--TKGELDLKLDLQASDVYVYRERLEQE 186
Oy 141 -----TTPFIPLMGNEKG-----MLTKKDAKOLEVAAQFTPLS 174
Db 187 KELLHNQNSMLNTELTAKTDBELLALGKREKNEILELCTLENNKKEEDAIRSHSSASASA 246
Oy 175 LSFPLDNNNPPLMSSRRPNRPVYLPPIFMHGKPNRSPNTPSHARQFTPNFRADLFKQV 234
Db 247 LS--SSPNLSTPGWSPKTP-----VPAQREPARSGSTOEKNNKIRPRGQDSSYYWEI 298
Oy 235 SVKKAEDLMGTDSDLMFGYTOOSHW-----QIFNGKNSRPRVHDYOPEI----- 281
Db 299 EASEVMLSTRIGSS---FGTVYKGAHGDVAVAKILKVVDPQLQAFRNEVAVLKRTR 355
Oy 282 -----FLTQ-----PVYSDL--PMDGKVRMI-----GMGAVNH 307
Db 356 HVNILLFMGYTKDNLAIIVQWCGSSLYKHLHYQETKRFQMFOLIDLARQAGQMDYLHA 415

```

```

Oy 308 SNGESAKLRSNMNRAYLMAGMEKN-----NLTPMPRIMGRIFKEGSGSO----- 351
Db 416 KNIIHRDKMS--NNIFLEGFLVKGIGDGLATVSRW-----SGSQVQDPPIGSVLM 466
Oy 352 -----PDDNPDLI--DYGYGDVRYLYOLENKSNIQTVARYNRRSGKALQLDYVPL 402
Db 467 APEVIRMQDNNPFSPQSDVYSYGIY--LYEL-----MTGELPYSHINNRDI-----IMV 515
Oy 403 GKG 405
Db 516 GRG 518

```

RESULT 15

```

ID 090N32 PRELIMINARY; PRT: 798 AA.
AC 090N32;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE PPAR GAMMA COACTIVATOR-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Larrouy D., Vidal H., Andreelli F., Laville M., Langin D.;
RT "Cloning and mRNA tissue distribution of human PPARgamma coactivator-
RT 1."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF159714; AAD51615.1; -
DR InterPro: IPR000504; RRM.
DR Pfam: PF00076; rrm.1.
DR PROSITE: PS50102; RRM.1.
DR SMART: SM00360; RRM.1.
SQ SEQUENCE 798 AA; 91056 MW; F68F9768BD94E1F4 CRC64;

```

Query Match 4.5%; Score 106; DB 4; Length 798;
 Best Local Similarity 24.1%; Pred. No. 1.5;
 Matches 93; Conservative 36; Mismatches 115; Indels 142; Gaps 27;

```

Oy 3 VSLSTLTLITLPCFAIILAIQQAQVPPVAVFVDEVRKNDLGONELLIGVQASSTQ 62
Db 89 LAVLETLDLSLPV-----DEGLPSFDALTD-----GDVYTTNE-----ASPSPM 129
Oy 63 D-TANPLDHEHEPELYTTAL--ENKTMLIN-CSALNODIMRLACYDTLVHGE-----T 110
Db 130 DGTPEPQAEHEEPLSKLLKLLAPANTOLSYNECSGLS-----TONHANHHIRITN 179
Oy 111 PAVIKTRKSIRIDETIMQT-----IKGKPQ-----VVOETPTD-----PIFLMGNE 151
Db 180 PAIKTENS-----WNNKAKSICQOQKQRRPQSELLKLTITNDPDPHTRKPPENRRSS 232
Oy 152 KGMLT--KKDAKOLEVAAQFTPLSFDLDNRNTPPLMSSRRPNMW-----VYLP 199
Db 233 RDKCTSKKSHSQSQSHLQAKPTLSL-----PLTPESPNDPKSGSPFNKTIERTLS 285
Oy 200 IFMNGKPNRS--PNTPSHARQFTPNFRRA--PELKQV-----SVKKAEDLMGT--- 247
Db 286 VELSGTAGLPPPTPHKRAQ--DNPFASPKLSSCKTVVPPSKKPRYSSESS--GTQGN 342
Oy 248 -----DSDLMFGYTOOSHWQIFNG-----KNSRP-----FRVHDY-----OPEIFL- 283
Db 343 NSTKKGPQSESL--YALSKSSSVLTGHEERKTRPBLRFLFGCHDYQGSINSTETILIN 399
Oy 284 -----TQPVYSDLPMDGKV 297
Db 400 ISOELQDSROLEKNRDVSSD--WGQGI 423

```

Search completed: November 30, 2001, 14:26:27
Job time: 564 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:16:59 ; Search time 72.04 Seconds
(Without alignments) 454.475 Million cell updates/sec

Title: US-09-787-083-6

Perfect score: 2360

Sequence: 1 MKVSLSTLTLSTLSCFALLA.....YVHAEATSEGVGLMDMGL 442

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_1101:*

1: /SIDSB/gcgdata/geneseq/geneseq/AA1980.DAT:*

2: /SIDSB/gcgdata/geneseq/geneseq/AA1981.DAT:*

3: /SIDSB/gcgdata/geneseq/geneseq/AA1982.DAT:*

4: /SIDSB/gcgdata/geneseq/geneseq/AA1983.DAT:*

5: /SIDSB/gcgdata/geneseq/geneseq/AA1984.DAT:*

6: /SIDSB/gcgdata/geneseq/geneseq/AA1985.DAT:*

7: /SIDSB/gcgdata/geneseq/geneseq/AA1986.DAT:*

8: /SIDSB/gcgdata/geneseq/geneseq/AA1987.DAT:*

9: /SIDSB/gcgdata/geneseq/geneseq/AA1988.DAT:*

10: /SIDSB/gcgdata/geneseq/geneseq/AA1989.DAT:*

11: /SIDSB/gcgdata/geneseq/geneseq/AA1990.DAT:*

12: /SIDSB/gcgdata/geneseq/geneseq/AA1991.DAT:*

13: /SIDSB/gcgdata/geneseq/geneseq/AA1992.DAT:*

14: /SIDSB/gcgdata/geneseq/geneseq/AA1993.DAT:*

15: /SIDSB/gcgdata/geneseq/geneseq/AA1994.DAT:*

16: /SIDSB/gcgdata/geneseq/geneseq/AA1995.DAT:*

17: /SIDSB/gcgdata/geneseq/geneseq/AA1996.DAT:*

18: /SIDSB/gcgdata/geneseq/geneseq/AA1997.DAT:*

19: /SIDSB/gcgdata/geneseq/geneseq/AA1998.DAT:*

20: /SIDSB/gcgdata/geneseq/geneseq/AA1999.DAT:*

21: /SIDSB/gcgdata/geneseq/geneseq/AA2000.DAT:*

22: /SIDSB/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2360	100.0	442	21	AA1985270
2	2352	99.7	442	21	AA1985268
3	2348	99.5	442	21	AA1985271
4	2330	98.7	442	21	AA1985269
5	822	34.8	370	21	AA1985156
6	815	34.5	370	21	AA1985157
7	815	34.5	370	21	AA1985158
8	815	34.5	374	21	AA1985159
9	812	34.4	375	21	AA1985160
10	246.5	10.4	355	19	AA1985161
11	239.5	10.1	356	19	AA1985162

12	156.5	6.6	253	18	AA1985270
13	114	4.8	1686	19	AA1985271
14	109.5	4.6	800	8	AA1985272
15	109.5	4.6	822	13	AA1985273
16	108	4.6	824	21	AA1985274
17	107	4.5	157	18	AA1985275
18	104.5	4.4	537	22	AA1985276
19	104	4.4	798	21	AA1985277
20	104	4.4	798	21	AA1985278
21	104	4.4	798	22	AA1985279
22	100.5	4.3	682	17	AA1985280
23	99.5	4.2	1726	18	AA1985281
24	98.5	4.2	467	20	AA1985282
25	98.5	4.2	467	21	AA1985283
26	98.5	4.2	467	22	AA1985284
27	98.5	4.2	467	22	AA1985285
28	95.5	4.0	502	22	AA1985286
29	95.5	4.0	516	22	AA1985287
30	95	4.0	522	22	AA1985288
31	94	4.0	372	16	AA1985289
32	94	4.0	761	20	AA1985290
33	94	4.0	1115	21	AA1985291
34	93.5	4.0	467	21	AA1985292
35	92.5	3.9	1717	22	AA1985293
36	92	3.9	888	22	AA1985294
37	91.5	3.9	857	13	AA1985295
38	91	3.9	1227	22	AA1985296
39	90.5	3.8	1024	18	AA1985297
40	90	3.8	460	21	AA1985298
41	90	3.8	481	21	AA1985299
42	90	3.8	522	22	AA1985300
43	90	3.8	564	21	AA1985301
44	90	3.8	564	21	AA1985302
45	90	3.8	600	22	AA1985303

ALIGNMENTS

RESULT 1	
AA1985270	standard; Protein: 442 AA.
ID	AA1985270
AC	AA1985270
XX	
DT	29-JUN-2000 (first entry)
XX	
DE	BASB034 amino acid sequence #3.
KW	Moraxella catarrhalis infection; BASB034: diagnosis; staging;
KW	vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW	sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW	hearing loss; antibacterial drug.
XX	
OS	Moraxella catarrhalis.
XX	
PN	W0200015802-AL.
XX	
PD	23-MAR-2000.
XX	
PF	14-SEP-1999; 99WO-EP06761.
XX	
PR	14-SEP-1998; 98GB-0020002.
PA	(SMIK) SMITHKLINE BECHAM BIOLOGICALS.
XX	
PI	Ruelle J;
DR	WPI: 2000-271440/23.
XX	
PT	N-PSDB; AAA10702.
XX	
PT	Novel BASB034 polynucleotides and polypeptides from Moraxella
XX	catarrhalis used to prepare vaccines against bacterial infections

H. pylori outer me
Human class II p13
Sequence encoded b
Alkaline cellulase
Bacillus sp. KSM-S
H. pylori outer me
C glutamicum prote
Human ORF689
Human PGC-1 protei
Peroxisome prolifer
S. pneumoniae peni
Phosphatidyl inosi
Amino acid sequenc
Human protein clon
Human membrane or
Human P80260 prote
C. glutamicum prote
Human protein sequ
Gonococcal porin-5
Non-B, non-C, non-
Amino acid sequenc
Human secreted pro
Ostrinia nubialis
S cerevisiae apopt
S receptor kinase
S. epidermidis ope
Mycoplasma genital
Arabidopsis thalia
Human protein sequ
T. malsutake pyran
Trichoderma derive
Human colon cancer

XX Claim 3; Page 68; 106pp; English.
PS
XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2913. The invention relates to BASB034 polypeptides from
CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
CC polynucleotides and polypeptides may be employed as research reagents and
CC material for the discovery of treatments and diagnostics for diseases,
CC particularly human diseases. They are particularly used to diagnose and
CC treat M. catarrhalis infections. They can be used for diagnosis of
CC disease, staging of disease, or determining response of an infectious
CC organism to drugs. The polynucleotides may be used as a source for
CC hybridization probes, and for screening of genetic mutations, serotype,
CC organism or strain identification, identification of mutations in BASB034
CC sequences, and as components of arrays which are useful for diagnostic
CC and prognostic purposes. The polypeptides can be used to produce
CC antibodies. The polypeptides can also be used in vaccine formulations,
CC and to identify agonists and antagonists. The polypeptides, antibodies,
CC agonists and antagonists (which are bacteriostatic) are used for the
CC treatment and prevention of diseases such as otitis media in infants and
CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
CC invasive diseases, and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.
CC
XX
SQ Sequence 442 AA;

Query Match 100.0%; Score 2360; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 1.3e-218;
Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSLSTLTSLSCFAILAIQAKAVPNVAFVDEVRSENDLGDONELPIDVOSATQSA 60
DB 1 mkvslstltslscfaiilaigqakavpnvatvdevrsendlgdndelridvgsatqsa 60
QY 61 STDANPLDEHEPELYTTALENKTMLINCALNODIMRLACYDTLVHGEPPAVYIKTRSI 120
DB 61 stdanpldehepeilyttalenktmlincalndimrlacydtlvngeppavlyktrsl 120
QY 121 RLDEETIMQTIKGRPOVYVOETTPFIPLMGNEKMLTKKDAKOLEYAKOFTPLSLSDLD 180
DB 121 rldeetimqtkgrpovvyvoettpfiflmgnkmltkkdakqleyaakqftplslsldld 180
QY 181 RNNPTLWSSRPNNPMYULPIFMHGKPRNSPNTSPSHEAROPTNFEFRAPETKFOYSVYKA 240
DB 181 rnnptlwssrphnmpmyulpifmngkprnsntpshearoptnfefrapetkfoysvyka 240
QY 241 AEDLWGTSDLWFGYTGQSHMQIFNGKNSRPFVNDYQPEIFLTQPYSDLPMDGKVYRM 300
DB 241 aedlwgtsdglwfygtqshmqifngknsrpfvndyqpelfltqpyssdlpmdgkvym 300
QY 301 GNGAVHHNSNGESAKLSNMRATILMAGMEKMLTVMPRIKGRIFKESSGSDPDNDPILD 360
DB 301 gngavhhsngesaklsnmratilmagmekmltvmprigrifkessgsqdpndnpild 360
QY 361 YYGVDVRFYLYOLENKSNIIGTVRYNRSKGALQLDYVYPLGKIGISYQIOFGYQOSL 420
DB 361 yygvdvrflylyolenksniigtvrynrsgkalqldyvyplygkigisyqiofygysl 420
QY 421 IDYNHEATSEFGVGLMDNMGL 442
DB 421 idynheatsfvgvglmndmngl 442

RESULT 2
ID AA85268 standard; Protein; 442 AA.
XX
AC AA85268;
XX

DT 29-JUN-2000 (first entry)
XX
XX BASB034 amino acid sequence #1.
DE
XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
XX hearing loss; antibacterial drug.
OS
XX Moraxella catarrhalis.
XX
XX WO200015802-A1.
XX
XX 23-MAR-2000.
XX
XX 14-SEP-1999; 99WO-EP06781.
XX
XX 14-SEP-1998; 98GB-0020002.
XX
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Ruelle J;
XX
XX WPI: 2000-271440/23.
XX
XX N-PSDB; AAA10700.
XX
XX Novel BASB034 polynucleotides and polypeptides from Moraxella
PT catarrhalis used to prepare vaccines against bacterial infections -
XX
XX
PS Claim 3; Fig 2; 106pp; English.

XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2931 (ATCC 43617). The invention relates to BASB034 polypeptides
CC from M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The
CC BASB034 polynucleotides and polypeptides may be employed as research
CC reagents and material for the discovery of treatments and diagnostics for
CC diseases, particularly human diseases. They are particularly used to
CC diagnose and treat M. catarrhalis infections. They can be used for
CC diagnosis of disease, staging of disease, or determining response of an
CC infectious organism to drugs. The polynucleotides may be used as a source
CC for hybridization probes, and for screening of genetic mutations,
CC serotype, organism or strain identification, identification of mutations
CC in BASB034 sequences, and as components of arrays which are useful for
CC diagnostic and prognostic purposes. The polypeptides can be used to
CC produce antibodies. The polypeptides can also be used in vaccine
CC formulations, and to identify agonists and antagonists. The polypeptides,
CC antibodies, agonists and antagonists (which are bacteriostatic) are used
CC for the treatment and prevention of diseases such as otitis media in
CC infants and children, pneumonia in elderlies, sinusitis, nosocomial
CC infections and invasive diseases, and chronic otitis media with hearing
CC loss. The polypeptides, agonists and antagonists are also used for
CC screening of antibacterial drugs. The BASB034 products of the invention
CC can be used screen for new antibacterial compounds that may target
CC resistant bacteria.
CC
XX
SQ Sequence 442 AA;

Query Match 99.7%; Score 2352; DB 21; Length 442;
Best Local Similarity 99.3%; Pred. No. 7.6e-218;
Matches 439; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSLSTLTSLSCFAILAIQAKAVPNVAFVDEVRSENDLGDONELPIDVOSATQSA 60
DB 1 mkvslstltslscfaiilaigqakavpnvatvdevrsendlgdndelridvgsatqsa 60
QY 61 STDANPLDEHEPELYTTALENKTMLINCALNODIMRLACYDTLVHGEPPAVYIKTRSI 120
DB 61 stdanpldehepeilyttalenktmlincalndimrlacydtlvngeppavlyktrsl 120
QY 121 RLDEETIMQTIKGRPOVYVOETTPFIPLMGNEKMLTKKDAKOLEYAKOFTPLSLSDLD 180
DB 121 rldeetimqtkgrpovvyvoettpfiflmgnkmltkkdakqleyaakqftplslsldld 180

QY 181 RNNPTLMSRRHNPMPYVLPITFMHGKPNRSPPTPSHEARQFPTNPNFRAPELKEOYSVKYKA 240
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elderly, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.

QY 181 RNNPTLMSRRHNPMPYVLPITFMHGKPNRSPPTPSHEARQFPTNPNFRAPELKEOYSVKYKA 240
 Db 181 rnnptlmsrrhnpmpyvlptfmgkpnrsptpsheaqfptnfnrappelkfyvsvkka 240
 QY 241 AEDLMGTDSDLMFGYTOOSHWOIFNGKNSRPRVHDYQPEIFLTPYVSDLPMDGKVMIM 300
 CC 241 aedlmgtddlmfgytqooshwoifngknsrprvhdyqpeifltpyvsdlpmdgkvmi 300
 Db 241 aedlmgtddlmfgytqooshwoifngknsrprvhdyqpeifltpyvsdlpmdgkvmi 300
 QY 301 GMGAVHSHNGESAKLSRSNRAIYLMAGMEKNTLTPMRIRIFRSGSGQPDNDPDL 360
 CC 301 gmgavhshngesaklsrsnraylmagmeknltvmpriwgrifksgsgqddndpdl 360
 Db 301 gmgavhshngesaklsrsnraylmagmeknltvmpriwgrifksgsgqddndpdl 360
 QY 361 YYGVDVRFVLQLENKNSISGTVRNPRSGKALQLDVYVPLGKISGFQIFQGYGSL 420
 CC 361 yygvdvrflyqlenknsisgtvrvnprsgkalgldvyplygkisygfqifgygsl 420
 Db 361 yygvdvrflyqlenknsisgtvrvnprsgkalgldvyplygkisygfqifgygsl 420
 QY 421 IDYNHEATSFVGLMLNDMGL 442
 CC 421 idynheatsfvgylmndmgl 442
 Db 421 idynheatsfvgylmndmgl 442

RESULT 3
 ID AAY85271 standard; Protein: 442 AA.
 XX AAY85271:
 XX 29-JUN-2000 (first entry)
 XX BASB034 amino acid sequence #4.
 XX
 KM Moraxella catarrhalis infection; BASB034: diagnosis; staging;
 KM vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KM sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KM hearing loss; antibacterial drug.
 XX
 OS Moraxella catarrhalis.
 XX
 PN W0200015802-A1.
 PD 23-MAR-2000.
 XX 14-SEP-1999; 99WO-EP06781.
 PF 14-SEP-1999; 98GB-0020002.
 PR (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 PA
 XX
 PI Ruelle J;
 XX
 XX WPI: 2000-271440/23.
 DR N-PSDB; AAA10703.
 XX
 PT Novel BASB034 polynucleotides and polypeptides from Moraxella
 PT catarrhalis used to prepare vaccines against bacterial infections
 XX
 PS Claim 3; Page 69; 106pp; English.
 XX
 CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 CC strain Mc2969. The invention relates to BASB034 polypeptides from
 CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2965. The BASB034
 CC polynucleotides and polypeptides may be employed as research reagents and
 CC material for the discovery of treatments and diagnostics for diseases,
 CC particularly human diseases. They are particularly used to diagnose and
 CC treat M. catarrhalis infections. They can be used for diagnosis of
 CC disease, staging of disease, or determining response of an infectious
 CC organism to drugs. The polynucleotides may be used as a source for
 CC hybridization probes, and for screening of genetic mutations, serotype,
 CC organism or strain identification, identification of mutations in BASB034
 CC sequences, and as components of arrays which are useful for diagnostic
 CC and prognostic purposes. The polypeptides can be used to produce
 CC antibodies. The polypeptides can also be used in vaccine formulations,

CC and to identify agonists and antagonists. The polypeptides, antibodies,
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elderly, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.

QY 1 MKVSLSTLTLSILSCFALIAIOAKAVPNPAFVDEVSENDLGQDNELPIDVOSATQSA 60
 Db 1 mkvslstltlsilscfalialioakavpnppafvdevrsendlgqdnelpidvqatqsa 60
 QY 61 SYDTANPLDEHEPELYTTALENKTMLINCASALNDIMRLACYDTLVHGETPAVITKRSI 120
 Db 61 stdtanpldehepeilyttalenktmlincsalngdimrlacydtlvhgetpaviktksi 120
 QY 121 RLDERIMQTIKGRQVYVQETDPIFLMGNEKGMILTKKDAOLEVAAQOFTPLSFDLD 180
 Db 121 rldetlmqtkgrqvyyvqetdprflmgnekgmiltkkdaolevaaqoftplsfdld 180
 QY 181 RNNPTLMSRRHNPMPYVLPITFMHGKPNRSPPTPSHEARQFPTNPNFRAPELKEOYSVKYKA 240
 Db 181 rnnptlmsrrhnpmpyvlptfmgkpnrsptpsheaqfptnfnrappelkfyvsvkka 240
 QY 241 AEDLMGTDSDLMFGYTOOSHWOIFNGKNSRPRVHDYQPEIFLTPYVSDLPMDGKVMIM 300
 Db 241 aedlmgtddlmfgytqooshwoifngknsrprvhdyqpeifltpyvsdlpmdgkvmi 300
 QY 301 GMGAVHSHNGESAKLSRSNRAIYLMAGMEKNTLTPMRIRIFRSGSGQPDNDPDL 360
 Db 301 gmgavhshngesaklsrsnraylmagmeknltvmpriwgrifksgsgqddndpdl 360
 QY 361 YYGVDVRFVLQLENKNSISGTVRNPRSGKALQLDVYVPLGKISGFQIFQGYGSL 420
 Db 361 yygvdvrflyqlenknsisgtvrvnprsgkalgldvyplygkisygfqifgygsl 420
 QY 421 IDYNHEATSFVGLMLNDMGL 442
 Db 421 idynheatsfvgylmndmgl 442

RESULT 4
 ID AAY85269 standard; Protein: 442 AA.
 XX AAY85269:
 XX 29-JUN-2000 (first entry)
 XX BASB034 amino acid sequence #2.
 XX
 KM Moraxella catarrhalis infection; BASB034: diagnosis; staging;
 KM vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KM sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KM hearing loss; antibacterial drug.
 XX
 OS Moraxella catarrhalis.
 XX
 PN W0200015802-A1.
 PD 23-MAR-2000.
 XX 14-SEP-1999; 99WO-EP06781.
 PF 14-SEP-1999; 99WO-EP06781.
 XX

Db	21	lqcalcdntvrltlaacydriflaaqlpsrsagqeggeskavnlletvrrssldkgeavilvek	80
Qy	141	TTDPFIPLMGNEKGLTKKDAKQLEVAKAQFTPLSLSEFDLRNN-TPRLSSRPHPMYLP	199
Db	81	ggdnl-----pdsagetadilytrpsltnyldikndlrglllgvrehnplymp	127
Qy	200	IFMHGKRNBSRPNRPSHEAR-QFTPNFGRAPRLKFGYSVKVKAEDLMGDSDFWFGYQQ	258
Db	128	fwynspnyapsrptqvtgtvteqekfsgqkrnaetklqvsfkslaeni fkradlwfygqr	187
Qy	259	SHMOIFN-GKNRSPFRFVNDQPEIFPTOPVYSDLPMDGKVRKIGAGVAHNSGSEAKLSR	317
Db	188	schwilyngqgktksapfrmtqykrelftqpykadlprgqrlmlngagfthnsgngsqstpestr	247
Qy	318	SWNRAYLMAEMENKLTVMPIRWIGRIFFKGGSGQPDNDLIDYYGQGVREFLYOLENKS	377
Db	248	swnllymagmewgkltlviprvwrafdq-sgdk-ndnpdiadmggygdvdklqyLndrq	305
Qy	378	NISSTVYVNRSRSGKGAQLDYYVPLCKISGTFQIFQIGSGSLIDYHNHATSFGYGLMLN	437
Db	306	nvysvlyltnpkltcygaleaaytflpkpklkvrvrfgfyggeslidylnhkqnglqglmfn	365
Qy	438	DMMGL 442	
Db	366	dwgdl 370	
RESULT	6		
AA75157			
ID	AA75157	standard; protein; 370 AA.	
XX	AA75157:		
XX	21-MAR-2000	(first entry)	
De	Neisseria meningitidis	ONF 582 protein sequence SEQ ID NO:1788.	
XX	Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;		
KW	antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;		
RV	antibacterial; gene therapy.		
XX	Neisseria meningitidis.		
XX	WO9957280-A2.		
XX	11-NOV-1999.		
XX	30-APR-1999;	99MO-US09346.	
XX	01-MAY-1998;	98US-0083758.	
PR	31-JUL-1998;	98US-0094869.	
PR	02-SEP-1998;	98US-0098994.	
PR	02-SEP-1998;	98US-0099062.	
PR	09-OCT-1998;	98US-0103749.	
PR	09-OCT-1998;	98US-0103794.	
PR	09-OCT-1998;	98US-0103796.	
PR	25-FEB-1999;	99US-0121528.	
XX	(CHIR) CHIRON CORP.		
PA	(GENO-) INST GENOMIC RES.		
XX	Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;		
PI	Petersen J, Plaza M, Rappoldi R, Ratti G, Scalato E, Scarselli M;		
PI	Tetzelin H, Venter JC;		
XX	WPI: 2000-062150/05.		
DR	N-PSDB: AA253919.		
XX	Novel Neisserial polypeptides predicted to be useful antigens for		
PT	vaccines and diagnostics		
XX	Claim 2; Page 904; 1453pp; English.		

Query Match	34.5%	Score 815	DB 21	Length 370				
Best Local Similarity	44.7%	Pred. No. 6.7e-70						
Matches 163	Conservative 67	Mismatches 111	Indels 24	Gaps 8				
CC	AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941							
CC	represent novel <i>Neisseria meningitidis</i> and <i>N. gonorrhoeae</i> polynucleotides							
CC	and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent							
CC	PCR primers used in the exemplification of the present invention. The							
CC	polypeptides, the polynucleotides, antibodies and compositions of							
CC	the invention can be used as vaccines, as diagnostic reagents, and as							
CC	immunogenic compositions. The polypeptides can be used in the							
CC	manufacture of medicaments for treating or preventing infection due to							
CC	<i>Neisseria</i> bacteria (e.g. meningitis and septicemia), to detect the							
CC	presence of <i>Neisseria</i> bacteria, or to raise antibodies. They may also							
CC	be used to screen for agonists or antagonists, which may themselves							
CC	have use as antibacterial agents. The polynucleotides of the invention							
CC	may also be used in gene therapy protocols.							
XX	Sequence 370 AA:							
YY	87 INCSALNODIMRLACDYDLVHGTEPAVI---KTKRSIRLDEETIMQTI-KGKPOVYVOE 140							
DB	21 lqcaaltatvtltaacydrifaaqldpsasgqegeskavlnltetrssldkgeavilvek 80							
YY	141 TTDPPIFLMGNEKGMFLKKDAKOLEYAKQFTPLSLSPFLDRNN-TPLWSSRPHNDYVLP 199							
DB	81 ggdaL-----padsagetadlytDlslmydldkndrlglllgvrehnpymlp 127							
YY	200 ITMHGKRPNSPTPTPSHEAR-QFTPMDFRRLKPFQYSVVKAEDLMGDSIDLMKGYNOQ 258							
DB	128 lwnpsnpyapaspfctgtvtgqekfsgqkkaekllqvstfskllaedi fktadlwfgytqr 187							
YY	259 SSMQJFFN-KGNSRPPRRVHDYOEIFLTOPVYDLDPMWDKVRMIGGAAVHNSGESAKSRL 317							
DB	188 schwqlynggrksapfncndykpelftqpvyadlpfgrtlmvgafvhnqsgngsgrprestr 247							
YY	318 SMNRAYLMAEMEMKMLTVMPRIWGRIFKEGSGSQDPDDNDLIDYVGYGDVRFYLOENKS 377							
DB	248 swrllyamagmewgkltvprvvrfafdq-sgdk-ndoppdiadmgdydvklygrLndrg 305							
YY	378 NISGVIRYRPRSGKALQLDYVRLPKGKISGTFQJLPOGIGOSLIDYNEHATSFVGLMLN 437							
DB	306 nvyavlyrnyprkfygaiaeaayffpikgkLkgvvrghfygsaeljdyhnhkqnglqigLmfn 365							
YY	438 DMWGL 442							
DB	366 dldgll 370							
RESULT 7								
AAY75158								
ID	AAY75158 standard: Protein; 370 AA.							
XX	AAY75158;							
XX	21-MAR-2000 (first entry)							
DE	Neisseria meningitidis ORF 582 protein sequence SEQ ID NO:1790.							
KW	Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;							
KM	antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;							
KM	antibacterial; gene therapy.							
XX	Neisseria meningitidis.							
OS	W09597280-A2.							
XX	11-NOV-1999.							
XX	30-APR-1999; 99WO-US09346.							

DB 310 nysvslrnpkkygaieaaytlpikgklkvvrgfthgyeslidyhnhkqnglqimfn 369
 OY 438 DWMGL 442
 DB 370 dldgl 374

RESULT 9

AA770628
 ID AAY70628 standard; Protein; 375 AA.

AC AAY70628;

DT 18-JUL-2000 (first entry)

DE Neisseria meningitidis serogroup B strain ATCC13090 BASB033 protein.

KW BASB033: diagnosis, prophylaxis: treatment; antibacterial; vaccine;

XX Neisseria meningitidis infection.

OS Neisseria meningitidis.

PN WO200015801-A1.

PD 23-MAR-2000.

PF 09-SEP-1999; 99WO-EP06718.

PR 14-SEP-1998; 98GB-0020003.

PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

PI Ruelle J;

DR MPI: 2000-271439/23.

DR N-PSDB: AA552133.

PT Isolated BASB033 polypeptides and polynucleotides of Neisseria meningitidis, useful for diagnosis, prophylaxis and treatment of N. meningitidis infection -

PS Claim 4; Page 58; 93pp; English.

CC The present sequence is a BASB033 protein from

CC Neisseria meningitidis serogroup B strain ATCC13090. The protein shows homology to the Klebsiella pneumoniae outer membrane

CC phospholipase A. The present sequence is useful for diagnosis, prophylaxis and treatment of N. meningitidis infection. It may also be

CC used for the discovery and development of antibacterial compounds and in vaccine compositions.

XX Sequence 375 AA;

Query Match 34.4%; Score 812; DB 21; Length 375;

Best Local Similarity 44.4%; Pred. No. 1,3e-69;

Matches 162; Conservative 68; Mismatches 111; Indels 24; Gaps 8;

OY 87 INGSALNODIMRLACYDLVHGETPAVI-----KTKRSIRLDETIMQTI-KGKPOVYVOE 140

DB 26 lqcaaltchvtivcydrifaaglpssagqegqeskavlnltetvssldkgaavlyvek 85

OY 141 TTDPIFMGNEKGMLTRKDAKOLEYAKOPTPLSLFSDLDRNN-TPLMSSRPHNPMYVLP 199

DB 86 ggal-----padsagetadlytptslmydlkdndlglytrehnpmylmp 132

OY 200 IFNHGKPNNSPNTPSHBAQ-QFTPNFRADELKFOYSVKYKAADLMWGTDSDLMFGYTQO 238

DB 133 lwnspnpyapsprtygtvqekfgqkkaeklkysfkakiaedlftktradiwfytyqr 192

OY 259 SHHQIIN-GKNSRPRFVNDORPILTORPYSLPMDGKVRMIGMGAHVHNSNESAKLSR 317

DB 193 sdkvqlyngqrksapfirtdykpeflfqpvykadlpfggtrlmqgafvthqsgnsgrpestr 252
 OY 318 SWNRATYLMAGEMKNLTVMPRIGRTFKEGSGOPDDNPILDYGGDVRFLYOLENKS 377
 DB 253 swnrlyamagmewgkltlvlprrvwarfdq-sgdk-ndqpdadymgygdvklqyrindrq 310
 OY 378 NISGTVRYNPRSGKALQLDYVYPLGKISGYEQIFQGYGSLIDYVHEATSPFGVGLMLN 437
 DB 311 nysvslrnpkkygaieaaytlpikgklkvvrgfthgyeslidyhnhkqnglqimfn 370

OY 438 DWMGL 442

DB 371 dldgl 375

RESULT 10

AAW98871
 ID AAW98871 standard; Protein; 355 AA.

AC AAW98871;

DT 31-MAR-1999 (first entry)

DE H. pylori GHPO 1723 protein.

KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;

XX peptic ulcer disease.

OS Helicobacter pylori.

PN WO9843478-A1.

PD 08-OCT-1998.

PF 01-APR-1998; 98WO-US06371.

PR 29-JUL-1997; 97US-0902615.

PR 01-APR-1997; 97US-083457.

PR 24-JUN-1997; 97US-0881227.

PA (HUMA-) HUMAN GENOME SCI INC.

PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.

PI Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;

DR MPI: 1998-542293/46.

DR N-PSDB: AAX14590.

PT New isolated Helicobacter polynucleotides - used to develop products for the diagnosis, prevention and treatment of Helicobacter

PT infections and gastrointestinal diseases

PS Claim 8; Page 1976-1977; 2054pp; English.

CC This sequence represents a Helicobacter pylori GHPO protein of the

CC invention. The polypeptides can be used for preventing or treating

CC Helicobacter infections, and gastroduodenal diseases associated with

CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be

CC used for the production of antibodies. The products can also be used for

CC detection and diagnosis.

XX Sequence 355 AA;

Query Match 10.4%; Score 246.5; DB 19; Length 355;

Best Local Similarity 25.1%; Pred. No. 3.7e-15;

Matches 82; Conservative 41; Mismatches 117; Indels 87; Gaps 11;

OY 157 KDKAKOLEYAKOFTPLSLFSDLDRNNTPLMSSRPHNPMYVLPITFMHGKPNRSPNTPSHE 216

DB 69 kkylnmdylygtvlfdrfthf-----tpdffgyhnpnlp----- 102

QY 217 ARQFTPNEFRAPELKFOYVSVKKAEDLMGTDSDLMFGYTOOSHQIENGKNSRPPFRVHD 276
Db 103 ---ygrnef-----kfqlsfvrvfthilwtkgcltlyatqtdwfglyndpqsapmrmn 154
QY 277 YQPEFLIQPVYSDLPWDGKV---RMIGGAVHNSNG--ESAKLSNSMNRNAVYLMAGMEMKN 332
Db 155 fmpelllyvypI-nfkpfqgklgnfseIwlgqhlngvgagcypfknk-----egnpdq 209
QY 333 LTVMPRI-----WGRIFKEGSGSQP-----DDNPD 357
Db 210 fpgqvnIvkdynqgkdvrwggcvsgqrpyvfrIwkekgjklmwaywpyrpydqsnpn 269
QY 338 ILDYGYGSDVRFly-----OLENKSNIsgfvrYRNPSRGALQLDVYVPLGKISGYF 410
Db 270 IldymygnakIdyrrgrhfhelqIdlydfqywyd--rwhgafrIgytIrnfpfyIya 327
QY 411 QIFOGYGOSLIDYNHEATSFVGLMLN 437
Db 328 qwfnngygdqIyeydvfnrIgyvgrIln 354

RESULT 11

AA10960
ID AA10960 standard; Protein: 356 AA.

XX AC AAY10960;
XX DT 08-JUN-1999 (first entry)

DE H. pylori ORF 07ap80601_5083193_f3_8 cell envelope protein.

KW Vaccine; probe; diagnostic; ORF; cell envelope protein;
KM secreted protein; cellular protein.

OS Helicobacter pylori.

PN W09818323-A1.

XX PD 07-MAY-1998.

XX PF 28-OCT-1997; 97WO-US19575.

XX PR 14-JUL-1997; 97US-0891928.

XX PR 28-OCT-1996; 96US-0739150.

XX PR 06-DEC-1996; 96US-0759739.

XX PA (ASTR) ASTRA AB.

XX PI Alm RA, Smith D;

XX DR WPI; 1998-271811/24.

XX DR N-PSDB; AAX30427.

XX PT Helicobacter pylori nucleic acids and proteins - used to develop

PS PT products for the detection, prevention and treatment of H. pylori

PS PS infections

PS PS Claims 27, 31; Page 160-161; 279pp; English.

XX CC Recombinant or substantially pure preparations of H. pylori polypeptides

XX CC are disclosed, together with the nucleic acids encoding them. In all,

XX CC 73 ORFs are shown. The proteins are variously cell envelope proteins,

XX CC secreted proteins or other cellular proteins. Vaccines containing the

XX CC nucleic acids or proteins are claimed, as are probes containing at least

XX CC 8 nucleotides from the nucleic acid sequences. The vaccines are useful

XX CC for treating or reducing the risk of H. pylori infections, and the

XX CC probes can be used diagnostically for detecting the presence of

XX CC Helicobacter in a sample. The products are also of use in screening

XX CC cycle or to inhibit H. pylori infection.

XX Sequence 356 AA;

Query Match 10.1%; Score 239.5; DB 19; Length 356;
Best Local Similarity 24.8%; Pred. No. 1.7e-14;
Matches 82; Conservative 43; Mismatches 111; Indels 95; Gaps 12;

QY 157 KDKAQLEIAAKQFTPLSFDLDRNNPFLMSSRRHPNRYVLPIMHGKPNRSPPTPSHE 216
Db 70 kKylmmmdylgyfIbIyhsf-----cpIlfqwyhpnIlnp----- 103
QY 217 ARQFTPNEFRAPELKFOYVSVKKAEDLMGTDSDLMFGYTOOSHQIENGKNSRPPFRVHD 276
Db 104 ---ygrnef-----kfqlsfvrvfthilwtkgcltlyatqtdwfglyndpqsapmrmn 155
QY 277 YQPEFLIQPVYSDLPWDGKV---RMIGGAVHNSNG--ESAKLSNSMNRNAVYLMAGMEMKN 332
Db 156 fmpelllyvypI-nfkpfqgklgnfseIwlgqhlngvgagcypfknk-----egnpdq 214
QY 323 -----YLMAGMEMK--NLTVPRIWGRIRKESGSGOPDD- 354
Db 215 pIvkrdngqkdvrwggcvsgxvxxgnIcflvIwkekgjklmwaywpyv-----pydq 266
QY 355 -NPDILDYGYGSDVRFly-----OLENKSNIsgfvrYRNPSRGALQLDVYVPLGKGI 406
Db 267 snpqIldymygnakIdyrrgrhfhelqIdlydfqywyd--rwhgafrIgytIrnfpfy 324
QY 407 SGYFQIFOGYGOSLIDYNHEATSFVGLMLN 437
Db 325 gIyagqwfnngygdqIyeydvfnrIgyvgrIln 355

RESULT 12

AAW20760
ID AAW20760 standard; Protein: 253 AA.

XX AC AAW20760;

XX DT 15-JUL-1997 (first entry)

DE H. pylori outer membrane protein, 07ap80601orf8.

KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;

KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;

KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope;

XX KM outer membrane.

XX OS Helicobacter pylori.

XX PN W09640893-A1.

XX PD 19-DEC-1996.

XX PF 06-JUN-1996; 96WO-US09122.

XX PR 01-APR-1996; 96US-0630405.

XX PR 07-JUN-1995; 95US-0487032.

XX PA (ASTR) ASTRA AB.

XX PI Berglindh OT, Smith D, Mellgaerd BL;

XX DR WPI; 1997-052306/05.

XX DR N-PSDB; AAT68013.

XX PT Helicobacter pylori nucleic acid sequences and related

XX PS PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori

XX PS PS infection, and to detect Helicobacter

XX PS Claim 56; Page 1172-1173; 1481pp; English.

XX CC The present sequence is a Helicobacter pylori outer membrane protein.

XX CC The protein may be used in a vaccine to prevent or treat H. pylori

XX CC infection or to identify H. pylori polypeptide binding compounds, useful

XX CC as potential H. pylori life cycle activators or inhibitors. The genomic

CC sequence of H. pylori (ATCC 55679) was determined from overlapping
CC contigs generated by mechanically shearing the bacterial DNA. The
CC sequences were analysed for ORF of at least 180 nucleotides, and the
CC predicted coding regions defined by computer evaluation. To identify
CC likely H. pylori antigens for vaccine development, the amino acid
CC sequences predicted from various ORF were analysed for significant
CC homology to other known or exported membrane proteins. Having identified
CC and determined the sequences of interest, particular regions can be
CC isolated from H. pylori by PCR amplification for recombinant polypeptide
CC production, e.g. in E. coli hosts.

CC
XX
SQ Sequence 253 AA:

Query Match 6.6%; Score 156.5; DB 18; Length 253;
Best Local Similarity 25.6%; Pred. No. 1e-06;
Matches 56; Conservative 26; Mismatches 72; Indels 65; Gaps 8;

OY 157 KKDAAQLAEAAQFPLSLSFOLDNRNTPLMSSRPINPMYVLPFIMHGKRNRPNSHE 216
DB 72 KKYIIMMDYIqYfIpfYhsf-----tpIqfWYhpnInp----- 105
OY 217 AAOFTFNERAPBELKQOVAVKAAEDLWGTDSLMFGYTOOSHOIIFNGKNSRPFVRHD 276
DB 106 ---ygrnef----kfqisfrvpvfrhllwtkgtlylaytqtnwfdqyndpgsapmrmn 157
OY 277 YQEPFLTPQVYSDLPMDCKV---RMIGGAVHHSNG-ESAKLSRSNMNRAVLMAGMEWKN 332
DB 158 fmpellyvypI-nfkpfqgkigfseIwlgwqhlinsvgagcyqpfh----- 204
OY 333 LFMVPRIMGRIFKESGSG--PDDNPDLIDYKGVDF 369
DB 205 -----kegnpenqfpgqpvIvKdyngqkdvrtw 231

RESULT 13

AAW70991
ID AAW70991 standard; protein; 1686 AA.

XX AAW70991;

DT 19-OCT-1998 (first entry)

XX Human class II p13 kinase-C2alpha.

XX Human, class II phosphoinositide lipid kinase; p13 kinase;

KM PIK3-C2alpha; Class II; resistance; Wortmannin; LY294002.

XX Homo sapiens.

OS WO9832864-A2.

PN 30-JUL-1998.

PD 27-JAN-1998; 98WO-GB00244.

XX 28-JAN-1997; 97GB-0001652.

PA (LUDW-) LUDWIG INST CANCER RES.

XX Domin J, Waterfield MD;

DR WPI; 1998-427960/36.

XX N-PSDB; AAV42920.

XX New nucleic acid encoding phosphoinositol kinase 3-C2 alpha or its
XX fragments - useful for, e.g. treatment of tumour cells where
XX phenotype is associated with expression of kinase

XX Claim 3; Fig 1; 52pp; English.

CC The present sequence represents a human class II phosphoinositide lipid
CC (p13) kinase designated PIK3-C2alpha. It is characterised as a class II

CC kinase due to the presence of a conserved C2 domain found in murine and
CC Drosophila class II p13 kinases, its apparent lack of a p85 binding site
CC and a substrate affinity to inositol lipids PtdIns and PtdIns(4)P. The
CC protein has resistance to p13 kinase inhibitors Wortmannin and LY294002.
CC Antibodies against the protein (optionally humanised), are used to
CC identify class II p13 kinases. Antisense sequences, antibodies or
CC dominant negative mutants of the p13-C2alpha protein, are useful in human
CC or veterinary medicine to block class II kinases. They can be used to
CC treat tumour cells where the phenotype is associated with expression of
CC p13-C2alpha protein.

CC
XX
SQ Sequence 1686 AA:

Query Match 4.8%; Score 114; DB 19; Length 1686;
Best Local Similarity 19.0%; Pred. No. 0.25;
Matches 85; Conservative 68; Mismatches 154; Indels 140; Gaps 21;

OY 45 QDNELPIDV-----QSNTOASSTDTANPLDE-----HEPELYTALENKTMLIN----- 8
DB 522 eddetpydlnkhlyqiekpckeamttrpveellidsynqvelalqlenqhravdqvkav 581
OY 89 ---CSALNODIMRLACYDTLVHGEPYAVIKTKRSIRDETIMQTIKGPQVYVQETDPI 145
DB 582 rkicsald-gvelal-----tesvklktravnlp-----ktadv 618
OY 146 FLMGNEKMLTKDAQALEAAQFPLSLSFOLDNRNTPLMSSRPINPMYV-----LP 199
DB 619 slfge-----dlrsst-rgslnpenpygvslnqltaa 651
OY 200 IF-----MHGKPNRPNSHEARQFTFNERAPBELKQOVAVKAAEDL---WGTSDDLW 252
DB 652 Iydllrhansgrpslccasssksvkewtltteqlgtlftf---aagisnswsnqky 707
OY 253 FGYYTOOSHOIIFNGKN-SRPF---RVADYQPEIFLTD-----PV-YSDLPMDGKVRMT 300
DB 708 yllcsish-----ngkdlfkpgskkvgtynkfnyllkwdelilfpqisqlplesvhl 763
OY 301 GMGAVHHSNGESAKLSRSNMN-----RAYLMAGMEWKNLYVMPRIWGRIFK 345
DB 764 lfglnqsgsgspdsnkqkrgpealgvslpdcfrfllcgv-----tkllylw----- 812
OY 346 EGSQSQPDNDPDLIDYGVGVDFVLYOLENKSNSICVFRVNPNSGKGLDLVDVYPLGKG 405
DB 813 --tsshtnsvpgvtckkgymerivlqvdfpspafdllyctpvavdsliqhmactend 870
OY 406 ISGVFOIFQGYGOSLIDYNHEATSGV 432
DB 871 ikg-----klldlhkdsisl 887

RESULT 14

AAW70420
ID AAW70420 standard; protein; 800 AA.

XX AAW70420;

DT 20-JAN-1991 (first entry)

XX Sequence encoded by cellulase gene derived from Bacillus sp. No. 1139.

XX Enzyme; cellotriose; cellotetrose; hydrolysis.

XX Bacillus sp. No. 1139.

OS

XX Key Location/Qualifiers
XX peptide 1..30
XX Protein 31..800

XX JP62232386-A.

XX 12-OCT-1987.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:17:45 ; Search time 36.79 Seconds
(Without alignments)
270.358 Million cell updates/sec

Title: US-09-787-083-6

Perfect score: 2360

Sequence: 1 MKVSLSTLTSLSCFAILA.....YNHEATSGVGLMNDMGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

212252

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100.5	4.3	682	3	US-08-481-435-6
2	100.5	4.3	1726	2	US-08-609-049A-30
3	100.5	4.3	1726	4	US-09-170-996-30
4	98.5	4.2	1658	2	US-08-609-049A-13
5	98.5	4.2	1658	4	US-09-170-996-13
6	90.5	3.8	1024	4	US-09-091-117-5
7	89.5	3.8	857	1	US-07-717-331E-2
8	88	3.7	537	2	US-08-633-879C-2
9	87.5	3.7	503	1	US-07-946-497-2
10	87.5	3.7	503	1	US-08-483-322-2
11	87.5	3.7	503	2	US-08-478-882-2
12	87.5	3.7	666	4	US-08-961-083-2
13	86	3.6	781	1	US-08-373-134D-2
14	86	3.6	781	2	US-09-114-637-2
15	86	3.6	1088	3	US-08-633-768A-1
16	85	3.6	816	1	US-07-731-157A-4
17	85	3.6	816	1	US-08-229-444B-2
18	85	3.6	816	2	US-08-541-780-4
19	85	3.6	985	5	PCT-US96-03916-6
20	85	3.6	985	5	PCT-US96-03916-66
21	84.5	3.6	774	3	US-08-802-632-2
22	84.5	3.6	774	3	US-09-073-354-1
23	84.5	3.6	774	3	US-08-656-005A-1
24	84.5	3.6	774	4	US-09-073-259-1
25	84.5	3.6	774	4	US-09-363-095-1
26	84.5	3.6	774	4	US-09-418-027-1
27	84	3.6	522	6	RE34606-6

28	84	3.6	1627	1	US-07-665-792E-9	Sequence 9, App1
29	83	3.5	320	2	US-08-245-511-4	Sequence 4, App1
30	83	3.5	320	2	US-08-600-993A-4	Sequence 4, App1
31	82.5	3.5	657	4	US-09-306-593-2	Sequence 2, App1
32	82.5	3.5	797	4	US-09-086-912-2	Sequence 2, App1
33	82	3.5	527	2	US-08-592-126-145	Sequence 145, App
34	82	3.5	527	2	US-08-687-080-48	Sequence 48, App1
35	82	3.5	535	2	US-08-633-879C-4	Sequence 4, App1
36	81.5	3.5	663	1	US-08-441-139-7	Sequence 7, App1
37	81.5	3.5	844	2	US-07-731-157A-6	Sequence 6, App1
38	81.5	3.5	844	2	US-08-541-780-6	Sequence 6, App1
39	81	3.4	355	1	US-07-946-497-5	Sequence 5, App1
40	81	3.4	355	2	US-08-483-322-5	Sequence 5, App1
41	81	3.4	355	2	US-08-478-882-5	Sequence 5, App1
42	81	3.4	1121	1	US-07-789-915A-2	Sequence 2, App1
43	81	3.4	1121	1	US-08-005-002C-2	Sequence 2, App1
44	81	3.4	1121	1	US-08-487-203A-2	Sequence 2, App1
45	80	3.4	334	2	US-08-359-850-4	Sequence 4, App1

ALIGNMENTS

RESULT 1
US-08-481-435-6
Sequence 6, Application US/08481435
Patent No. 6027906
GENERAL INFORMATION:
APPLICANT: Balganesht, Tanjore S
TITLE OF INVENTION: No. 6027906el Polypeptides
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESSES:
ADDRESSEE: White & Case
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,435
FILING DATE: 10-JUL-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IN 580/MAS/94
FILING DATE: 01-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9404072-2
FILING DATE: 24-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Steiner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-151
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 682 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-481-435-6

Query Match 4.3% Score 100.5; DB 3; Length 682;
Best Local Similarity 20.4%; Pred. No. 0.11;
Matches 92; Conservative 68; Mismatches 189; Indels 103; Gaps 24;

ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-0637000S
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ. ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 1726 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-170-996-30

Query Match 4.3%; Score 100.5; DB 4; Length 1726;
Best Local Similarity 18.6%; Pred. No. 0.54;
Matches 98; Conservative 79; Mismatches 190; Indels 161; Gaps 21;

QY 1 MKVSL-----TLTSLSCFAIIAIOAKAVNPVAFV-----EYRSENDLGOD 46
DB 465 VKVSIIEGLQLPVTFCDVSTVEIIMQALSWHDLQVDVGYTLKVGGEVLQ 524
QY 47 NELPIDVQATQASPTDANPDEHEPELTTALENKTMLINCSALMODIMRLACYDTLV 106
DB 525 NHCIGSHEHIONCKRMDT-----ETKQLTLTUSAMCONIARTAEED--- 565
QY 107 HGETPAVITKRSIRIDETIMOTIKGPKQVVOETTPIF----- 146
DB 566 --EAP-----VDLNKYLQIEKPYKEVMTRHPVELLDSDHYQVELALQTFNQHRV 615
QY 147 -----LMGNEKMLTKKDAKOLEYAA-----KQTPPLSLSDLDLNNTPLWS 188
DB 616 DOYIKAVRKICSLADGVEPTVTEA--VKKLKRAVNLPRKNSADVTSLSGSDTRKNSYKGS 674
QY 189 SRPHNPVYLP-----IFMHGKPR-----SPNPSHEARQTFPNPEFAELKFOY 234
DB 675 LNPENPVQVSMHLTRITRYDLRLHANSSRCSTGCPGRSRNIKEAWTATE-----QL 726
QY 235 SVKKAEDL---WGTDSDLMPGYTOOSHMOIFNGKN--SRPF---RVHDYQPEIFLQ-- 285
DB 727 QLFVYAAHGSISSNMVSNVEKYLLCSLSH---NGKDLFRPIQSKKGTGYNKPFYLLKMD 782
QY 286 ----PV-YSDLPMDGKVMIGMGAHVHNSGESAKLSRSWN-----RAYL 324
DB 783 ELIIFPIQISQLPLESVLHLLTFGLVNLNOSGSSPDSNKKORKEPALGVSLTLDFKRL 842
QY 325 MAGMEKNLVMRIRIRIKESGSPDNDPILDYGGDVRFYQLENKSNISGTVR 384
DB 843 TCG-----TKLLYLW-----TSSHTNSIPGAIIPKSYVMERIVLYQVDFPSPADIT 889
QY 385 YNPRSGKALQLDVYVPLGKIGSYFOIFOGYGOSLIDYVNHGATSFV 432
DB 890 TSPQIDRNITIQDKLETLESIDKG-----KLLDIHRDSSFGL 927

RESULT 4
US-08-609-049A-13
Sequence 13, Application US/08609049A
Patent No. 5948664
GENERAL INFORMATION:
APPLICANT: Williams, Lewis T.
APPLICANT: Moliz, Lisa
APPLICANT: Chen, Yen-Wen
TITLE OF INVENTION: No. 5948664el PI 3-Kinase Polypeptides
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA

ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/609,049A
FILING DATE: 29-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-0637000S
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ. ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1658 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-609-049A-13

Query Match 4.2%; Score 98.5; DB 2; Length 1658;
Best Local Similarity 18.9%; Pred. No. 0.81;
Matches 101; Conservative 78; Mismatches 195; Indels 159; Gaps 23;

QY 5 LSTLTSLSCFAIIAIO--QAKAVNPVAFVDEVS-----ENDLGODNE 48
DB 381 LSPVTYVRNMGENASVKSIEIGQLPVTFCDVSTVEIIMQALCWVHDLNQ-- 437
QY 49 LPIDVOS-----ATQASPTDANPDEHE--PELTTALENKTMLINCSALNODIMRLAC 101
DB 438 --VDVGSYILKVGGEVYLNHCLSGHEHIONCKRMDTEKQLTLTUSAMCONIARTAE 495
QY 102 YDTLVHGETPAVITKRSIRIDETIMOTIKGPKQVVOETTPIF----- 146
DB 496 DD-----EAP-----VDLNKYLQIEKPYKEVMIRHPVELLDSDHYQVELALQTFN 542
QY 147 -----LMGNEKMLTKKDAKOLEYAA-----KQTPPLSLSDLDLNNTPLWS 183
DB 543 QHRAVDVIRAVRKICSLADGVEPTVTEA--VKKLKRAVNLPRKNSADVTSLSGSDTRK 601
QY 184 TPLMSSRPHNPVYLP-----IFMHGKPR-----SPNPSHEARQTFPNPEFAPE 229
DB 602 STKGLNPNPNVQVSMHLTTATITDRLHANSSRCSTGCPGRSRNIKEAWTATE----- 656
QY 230 LKQVSVKVAEAEDL---WGTDSDLMPGYTOOSHMOIFNGKN--SRPF---RVHDYQPEIF 282
DB 657 ---QLQFTVYAAHGSISSNMVSNVEKYLLCSLSH---NGKDLFRPIQSKKGTGYNKPFY 709
QY 283 LTO-----PV-YSDLPMDGKVMIGMGAHVHNSGESAKLSRSWN----- 320
DB 710 LKMDLLEIPIQISQLPLESVLHLLTFGLVNLNOSGSSPDSNKKORKEPALGVSLTLDF 769
QY 321 -RAYIMAGMEKNLVMRIRIRIKESGSPDNDPILDYGGDVRFYQLENKSNISGTVR 379
DB 770 FKRLTGC-----TKLLYLW-----TSSHTNSIPGAIIPKSYVMERIVLYQVDFPSPA 816
QY 380 SGTVYVNPBSGKALQLDVYVPLGKIGSYFOIFOGYGOSLIDYVNHGATSFV 432
DB 817 FDIYTSQIDRNITIQDKLETLESIDKG-----KLLDIHRDSSFGL 859

RESULT 5
US-09-170-996-13
Sequence 13, Application US/09170996
Patent No. 6291220
GENERAL INFORMATION:

APPLICANT: Williams, Lewis T.
 APPLICANT: Moiz, Lisa
 APPLICANT: Chen, Yen-Wen
 TITLE OF INVENTION: No. 6291220e1 PI 3-Kinase Polypeptides
 NUMBER OF SEQUENCES: 32
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, 8th Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-3834
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/170,996
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/609,049
 FILING DATE: 29-FEB-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Dow, Karen B.
 REGISTRATION NUMBER: 29,684
 REFERENCE/DOCKET NUMBER: 2307K-063700US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-326-2400
 TELEFAX: 415-326-2422
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1658 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-170-996-13

Query Match 4.2%; Score 98.5; DB 4; Length 1658;
 Best Local Similarity 18.9%; Pred. No. 0.81;
 Matches 101; Conservative 78; Mismatches 195; Indels 159; Gaps 23;

QY 5 LSTLTSLSCFAILAIQ--QAKAVNPFAVDEVRS-----ENDLGQDNE 48
 DB 381 LSPVTVQRMCGEMASVKYSIEIGLQLPVFTCDVSSIVEIIMQALCWMVDDLQ---- 437
 QY 49 LPIDVOS-----ATOSASTDTNPDLDEHE--PELYTTALENKTMLINCALNODIMRLAC 101
 DB 438 --VDVGSYILKVGQGEVLQNNHCISHEHIONCRKMDPEIKIQLTLTSLAMQNLARTAE 495
 QY 102 VDTLVHGETPAVIKTKRSIRLDETIMQTIKGRPOVVOYQETTPRIF----- 146
 DB 496 DD-----EAP-----VDLKKYILQLEKPYKEVMIRHPVEELLSYHVOVELAQTEN 542
 QY 147 -----LMGNEKGLTKKDAKOLEYAA-----KQFTPLSTSFDDIRNN 183
 DB 543 QHRAVDQVIKAVRKICSAIDGVETPSVTEA--VKKILRAVNLRRNSADVTISLGSSTRKN 601
 QY 144 TPLMSSRPINPMTVLP-----IFMHGKPNR---SPMTPSHARQFTPNNEFRAD 229
 DB 602 STKSLGLENPNPQVSMIDLTAIYDLRLHANSSRCSTGCPKRSRIKEMATATE----- 656
 QY 230 LKFOYSVKYKAEDL--WGTOSDLMEFGTQOSHMOIFNGKN--SRF---RVHDIQPELF 282
 DB 657 ---QLOFTVYAAHGSISSNNVSNVEKYLLCSLSH---NGKDLFKIQSKKVGTYKNFEY 709
 QY 283 LTVQ-----PV-YSDLPMDGKVRMIGMGAHVHNSGESAKLRSRN----- 320
 DB 710 LKIMDELIIFPIQIOLPLESVILHLTLFGVLANOSSSSSDSKOKRGPRALGKVSLLTLD 769

QY 321 -RAYLMAGMEMKNLVPMRIMGRIFKESSGSDPDNDILDYGYGVDFVLEKNSNI 379
 DB 770 KRFRLCG-----TKLILW-----TSHTNSIPGAIIPKRSYMERIVYQVDFPSPA 816
 QY 380 SGTAVRNPSSGALQLDVYVPLGKISGYFQIFQYGGSLDYNHEARFSGV 432
 DB 817 FDIYTSPOIDRNIIOODKLETLESIDIK-----KLIDIHRSSFFL 859

RESULT 6
 US-09-091-117-5
 Sequence 5, Application US/09091117
 Patent No. 6171589
 GENERAL INFORMATION:
 APPLICANT: The University of Melbourne
 TITLE OF INVENTION: Mycoplasma Recombinant Polypeptides and
 TITLE OF INVENTION: Vaccines
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GREENLEE, WINNER and SULLIVAN P.C.
 STREET: 5370 Manhattan Circle, Suite 201
 CITY: Boulder
 STATE: Colorado
 COUNTRY: United States of America
 ZIP: 80503
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/091,117
 FILING DATE: 12 JUNE 1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/AU96/00803
 FILING DATE: 13-DEC-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: AU PN7127
 FILING DATE: 13-DEC-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: WINNER, Ellen P.
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: +1 303 499 8080
 TELEFAX: +1 303 499 8089
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1024 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: Mycoplasma genitalium
 US-09-091-117-5

Query Match 3.8%; Score 90.5; DB 4; Length 1024;
 Best Local Similarity 20.2%; Pred. No. 2.6;
 Matches 86; Conservative 61; Mismatches 172; Indels 107; Gaps 22;

QY 1 MKVS-LSTLTSLSCFAILAIQQAQAVNPFAVDEVSENDLGQDNEPLPI-DVQSAHQ 58
 DB 1 MKISTTTTCLISGAFGTTAI---ALPTVALLNHOQONTKQON--PIKIDRFGIN 54
 QY 59 SASDTANPLDEHEPELTYTALENKTMILNCALNODIMRLACYDPLVHGETPAVIKTKR 118
 DB 55 NVOVPMTPL--HQYVEVT--NNKAIYDKAPQKFFL-----AKSALNNKL 98
 QY 119 SIRLDETIMQTIKGRPOVVOYQETTPDIFLMGNEKGLTKKDAKOLEYAKO--FTPLSLSP 177
 DB 99 QVEFDFILRT-----GVINLNLMDLKKWIDQTLFIPQSF 135
 QY 178 DLDNRNTPLMSSRPINPMTVLP--IFMHGKPNRSP--NTPSHARQFTPNNEFRABELKFOYS 235

```

Db 136 DLSANKLNTLNSQSEVSLDLEFIFTFNSDPKNOPLKLPDGSVYVANE-----SYTSYK 191
QY 236 VKAKAEDLMGTSDLMFQTOQSHWOIFNGKRSRPRVHDYQPEI-FLTQPIYSDL-PM 293
Db 192 ATLQKTKVLYTSHADSHVSGISYAIPTVSLNGKQNDSEFNPFSKNINFARKNYNMALNPF 291
QY 294 DGRVIRIGCAVH-----NGESAK-----LSRSMNRAVYLMAGHEMKNLVMPRIWGRIFK 345
Db 252 EAOQVYVGQKFLNOKVYNADVDKNDINNHIETQFNVAKITA-----TLIGKAFK 300
QY 346 ---EGSGSOP-----DDNPDLIDYV--GYGDV-----RELYOL-- 373
Db 301 QCEHKNQGPLSLKVLKVLGSLNNEFKQLFNVRGLDGFVSDLIQSSQSSNKKTVQDLF 360
QY 374 EKRSNI 379
Db 361 ENKTTI 366

```

RESULT 7

```

US-07-717-331F-2
; Sequence 2, Application US/0717331F
; Patent No. 5484905
; GENERAL INFORMATION:
; APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua
; APPLICANT: Stein
; TITLE OF INVENTION: A Receptor Protein Kinase Gene
; TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/717,331F
; FILING DATE: June 19th 1991
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 857 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-717-331F-2

```

Query Match 3.8%; Score 89.5; DB 1; Length 857;

Best Local Similarity 21.4%; Pred. No. 2.4; Mismatches 110; Indels 119; Gaps 23;

```

Matches 76; Conservative 50; Mismatches 110; Indels 119; Gaps 23;
QY 105 LVHGETPAVYIKTKRSIRLDETIWQITKGRPOVVOYQETDPIFLMG-----NEK---GLMTR 157
Db 24 LIHPALSIYINT-----LSSTESLITSSNKTLY---SPGSIFFVGFRITSRYVLGWMYK 75
QY 158 K-DAKOLEVAAKQFTPLSLD-----LDNRNTPLMSSRPHNPVYLPITMHC 205
Db 76 KVSDRYVYVAMNDPLSNAGITLKISGNLVLDHSNRPVMTN-----LTRGN 125
QY 206 PNRSP-----NTPSHEARQFT-----PNEFRAPELKFOYSVKAKAEDLM 245

```

```

Db 126 -ERSPVYAEELLANGNFVMRSSNNDASEYLMQSRDYDTDLLPEMKGYIKT----- 177
QY 246 GTDSDLMFQYTOQSHWOIFNGKRSRPRVHDYQ-----PEILTO---PVYSDLPMDGK 296
Db 178 GLNRFL-----TSMRSSDPSGNNF---SYKLETSQLEPEFLSRBNPFMHRSGPWNG- 226
QY 297 VRMIGKAVH-----HSNGESAKLSRSNRAI-----LMAGHEMKNLVMP--RI 339
Db 227 IRRSGIPEQKLSYMYNFIENNEEVAVTFMTNNSFSRLTSLISEGFQRLTWYPSIRI 286
QY 340 WGRIFKEGSGSOPD-----NPDLIDYGYGDVREPLYOLEKNSNIGTVR-YNPRS 389
Db 287 WNRFWSSPVDRQCDTIYMGCP-----YAYCDV-----NTSPVCNCIOGFNPRN 329

```

RESULT 8

```

US-08-633-879C-2
; Sequence 2, Application US/08633879C
; Patent No. 5928922
; GENERAL INFORMATION:
; APPLICANT: Kivirikko, Kari I.
; APPLICANT: Pihlajaniemi, Taina
; APPLICANT: Heilaakoski, Tarja I.
; APPLICANT: Annunen, Pia P.
; APPLICANT: Nissi, Ritva K.
; APPLICANT: No. 5928922Jelainen, Minna K.
; TITLE OF INVENTION: 2 SUBUNIT OF PROLYL-4-HYDROXYLASE
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING SUCH SUBUNIT AND
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/633,879C
; FILING DATE: 10-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Abrams, Samuel B
; REGISTRATION NUMBER: 30,605
; REFERENCE/DOCKET NUMBER: 8389-0041-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 537 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-633-879C-2

```

Query Match 3.7%; Score 88; DB 2; Length 537;

Best Local Similarity 21.4%; Pred. No. 1.6; Mismatches 100; Conservative 52; Mismatches 145; Indels 170; Gaps 26;

TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-483-322-2

Query Match 3.7%; Score 87.5; DB 1; Length 503;
Best Local Similarity 18.4%; Pred. No. 1.6;
Matches 57; Conservative 31; Mismatches 97; Indels 125; Gaps 14;

QY 40 ENDIGQDNELPIDVQASATDTNPANLDEHEBELYTTALENTMLNCSALMODIMRL 99
DB 169 QEDIDASNIIDEVSS-----GSTIEKSTPEGYILHTDLP-----SQTGDRD----- 212
QY 100 ACYDTLVHGETPRAVIKTKRSI-----RLDETIWQTIKGRQVYVYQETDPIFLMGNEKG 153
DB 213 ---DAFPGSTLATIATTPWVSAHTKONQERTOWNPIHSNPEVLLQTT----- 258
QY 154 MLTKKDAKQLEYAKQFTPLSLSFIDLRRNTPL-----WSSRPHNPMVVLPIFMHGRPNRS 209
DB 259 -----RMT-----DIDRNSTSAHGENMTQEPQPP----- 283
QY 210 PNTPSHARQFTPNFPAPELKFQVSVKVAEDLW-----GTSDSLMGYTTQOS 259
DB 284 -NNHEYODEEETPH-----ATSTTWADPNSTTEAATQKEKMF-----EN 322
QY 260 HMOIFNGKN-----SRFRVHDYQPELFILOPYVSDLPW-----DGKVRMIMG 303
DB 323 EMQ---GKNPPTPSSEDSHVTEGTTASAHNNHPSQRMPTQSOEDVSWTDFDPISHPMGOG 379
QY 304 AVHHSNGESA 313
DB 380 HQTESKGHSS 389

RESULT 11
US-08-478-882-2
; Sequence 2, Application US/08478882
; Patent No. 5885575
; GENERAL INFORMATION:
; APPLICANT: HERRLICH, Peter
; APPLICANT: PONTA, Helmut
; APPLICANT: GUENTHERT, Ursula
; APPLICANT: MATZKU, Siegfried
; APPLICANT: WENZL, Achim
; TITLE OF INVENTION: VARIANT CD44 SURFACE PROTEINS, DNA
; TITLE OF INVENTION: SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,882
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/946,497
; FILING DATE: 19921109
; ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16915/145
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-478-882-2

Query Match 3.7%; Score 87.5; DB 2; Length 503;
Best Local Similarity 18.4%; Pred. No. 1.6;
Matches 57; Conservative 31; Mismatches 97; Indels 125; Gaps 14;

QY 40 ENDIGQDNELPIDVQASATDTNPANLDEHEBELYTTALENTMLNCSALMODIMRL 99
DB 169 QEDIDASNIIDEVSS-----GSTIEKSTPEGYILHTDLP-----SQTGDRD----- 212
QY 100 ACYDTLVHGETPRAVIKTKRSI-----RLDETIWQTIKGRQVYVYQETDPIFLMGNEKG 153
DB 213 ---DAFPGSTLATIATTPWVSAHTKONQERTOWNPIHSNPEVLLQTT----- 258
QY 154 MLTKKDAKQLEYAKQFTPLSLSFIDLRRNTPL-----WSSRPHNPMVVLPIFMHGRPNRS 209
DB 259 -----RMT-----DIDRNSTSAHGENMTQEPQPP----- 283
QY 210 PNTPSHARQFTPNFPAPELKFQVSVKVAEDLW-----GTSDSLMGYTTQOS 259
DB 284 -NNHEYODEEETPH-----ATSTTWADPNSTTEAATQKEKMF-----EN 322
QY 260 HMOIFNGKN-----SRFRVHDYQPELFILOPYVSDLPW-----DGKVRMIMG 303
DB 323 EMQ---GKNPPTPSSEDSHVTEGTTASAHNNHPSQRMPTQSOEDVSWTDFDPISHPMGOG 379
QY 304 AVHHSNGESA 313
DB 380 HQTESKGHSS 389

RESULT 12
US-08-961-083-2
; Sequence 2, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders
 REGISTRATION NUMBER: 36,373
 REFERENCE/DOCKET NUMBER: P8340P2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 666 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-961-083-2

Query Match 3.7%; Score 87.5; DB 4; Length 666;
 Best Local Similarity 20.1%; Pred. No. 2.6;
 Matches 88; Conservative 64; Mismatches 188; Indels 97; Gaps 24;

25 KAVNPVAVFVSEVSEND-GODNELPIDVQSATOS-----ASTDPANPLDHEHE 74
 1 KIYDKNQIADIGSERRRVNAQANDIPTLVKAIYSIEDHRRFDHNGIDTIRILGAFKRN 60
 75 LVTTALENKTMLNCSALNODIMRLACYDYLHGETPAVYIKTKRSIRLDETIWQTIKGR 134
 61 LQSNISLOGS-----SLTQQLIKLYFSTSDQTI-----RKAQEAFLAIQLDQ 106
 135 QVVYQET-----IDPFLMENEKGMILK-----KDAKOLE-----YAAKFPTL 173
 107 KATKQILTYINKYVMSGNGTGAQNYGKDLNLSLQALLAAGPQAPNQYDYL 166
 174 S-LSEFDLNNPTPLMSSRPNNPMPYLPFMHGKPNRSPNTPSHARQFTPNFEHRAPE--- 229
 167 SHPEAQAQDRNLVLT--SEKKNQGYI-----SAGEYKAVVTPTLDGQSLKSNAYAYMD 220
 230 --LKFOVS--VKYKAEDLMGTSDDLMEGYTQ--QSH-WOIFNGKNSRPVRHVDYQEPFL 283
 221 NTLKEVINQVEETGNTLLTGTGMDVYTNVDOEAKHLMIDYTNDEVAVYAPDDELQ----- 275
 284 TQPVYSDDLMDPKVMMIGMAVHNSN-----GESAKLSRNRATLVMGMEKNLT-VM 336
 276 VASTIYDVS-NGKV-LAQGARHQSNNVSFGINQAVETNRDW-----GSTMKPTDYA 326
 337 PRI-----WGRIFKESGSOQDDNPDLIDY-YGY-GDVRFYQLENKSNISGTVRYN- 386
 327 PALEYGVYSTATYVDEYNYNGTNTPYVNMDRGYFGNITLQYALQSRNPVAVETLTK 386
 387 -----PRSGKALQLDY 398
 387 VGLNRAKTFNLNGIDY 403

RESULT 13
 US-08-373-134D-2
 Sequence 2, Application US/08373134D
 Patent No. 5780296
 GENERAL INFORMATION:
 APPLICANT: Kmelec, Eric
 APPLICANT: Hollioman, William
 TITLE OF INVENTION: COMPOSITIONS AND METHODS TO PROMOTE
 TITLE OF INVENTION: HOMOLOGOUS RECOMBINATION IN EUKARYOTIC CELLS AND ORGANISMS
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/373,134D
 FILING DATE: January 17, 1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Friedel, Thomas E.
 REGISTRATION NUMBER: 29,258
 REFERENCE/DOCKET NUMBER: 7991-007
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-9741/8864
 TELEX: 66141 PENNITE
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 781 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-373-134D-2

Query Match 3.6%; Score 86; DB 1; Length 781;
 Best Local Similarity 20.7%; Pred. No. 4.9;
 Matches 50; Conservative 37; Mismatches 122; Indels 32; Gaps 9;

39 SENDLGODNELPIDVQS-ATQASDTPLANPLDHEHELYTTALENKTMLNCSALNODIM 97
 533 ARNDPQTSSQAPTSASFSGTGSALPDQPLAMDVASQATFTSGLIASIPATLAEVAREI 592
 98 RLACYDYLHGETPAVYIKTKRSIRLDETIWQTI-----KGRPOVYQET-----TTDPI 145
 593 DSAC-----ASNDVPLRLTLEARTAQLOGT--WSNLINRVFLSKTRARICRRDDQAPACEV 647
 146 FLMGNEKGMILTKDAQOLEYAKQFTPLSLF-FDLDRNNTPLMSSR-----PHNPMYVLEI 200
 648 RQNTNORGTASKSLMNTVTKAAVAVINPFGATYMLDVGVDSKSLQLRFRVITPRKAVHL-- 705
 201 FMHGKPNRSPNTPSHARQFTPNFEHRAPELVKFOVSVKYKAEDLMGTSDDLMEGYTQOSH 260
 706 -----NATPSTYVM-AMNATDSTPAESQOQRAAERHPAQEDADDLREGALQENH 758
 261 W 261
 759 W 759

RESULT 14
 US-09-114-637-2
 Sequence 2, Application US/09114637
 Patent No. 5945339
 GENERAL INFORMATION:
 APPLICANT: Kmelec, Eric
 APPLICANT: Hollioman, William
 TITLE OF INVENTION: COMPOSITIONS AND METHODS TO PROMOTE
 TITLE OF INVENTION: HOMOLOGOUS RECOMBINATION IN EUKARYOTIC CELLS AND ORGANISMS
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/114,637
 FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/373,134
FILING DATE: January 17, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Friedel, Thomas E.
REGISTRATION NUMBER: 29,258
REFERENCE/DOCKET NUMBER: 7991-007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 781 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-114-637-2

Query Match 3.6%; Score 86; DB 2; Length 781;
Best Local Similarity 20.7%; Pred. No. 4.9;
Matches 50; Conservative 37; Mismatches 122; Indels 32; Gaps 9;

QY 39 SENDLDQDNELPIDVOS-ATQASSTDTANPLDHEPELTYTALENKTMILNCSALNODIM 97
DB 533 ARNDPPTSOAPTSAFSGGSGALPDOPPLAMDVAQSOTAFISGLIASIAPTLAEAVGAREL 592
QY 98 RLACYDTLVHGETPAVYIKTKRSIRLDETIMQTI-----KGKPOVYOE---TTDPI 145
DB 593 DSAC-----ASNDVPLRLTKLKTALQLOTF-WSNLINVRFLSKTRARICMRDDOAPACEPV 647
QY 146 FLMGNEKMLTKKDAKOLEYAAKQFTPLSL-EDLDRNNTPLWSSR---PHNPMVYLP 200
DB 648 RQTNORGATSKSLMTMVRKAAVIVNPGATMLDVGVDKSGALQLRQVITPRKAVHVL- 705
QY 201 FMHGKRNRSNTPSHARQFTNEFRAPRLKFOYSKVAKAEDLMGDSLMTGYTQOSH 260
DB 706 -----NAYPTVYM-AMHATADSTPAPESOQOORAAERHPABOEDADDLFGELQEH 758
QY 261 W 261
DB 759 W 759

RESULT 15
US-08-633-768A-1
Sequence 1, Application US/08633768A
Patent No. 6013504
GENERAL INFORMATION:
APPLICANT: YU, SHUKUN
APPLICANT: BOJSEN, KIRSTEN
APPLICANT: KRAGH, KARSTEN
APPLICANT: BOJKO, MAJA
APPLICANT: NIELSEN, JOHN
APPLICANT: MARCUSSEN, JAN
TITLE OF INVENTION: ALPHA-1,4-GLUCAN LYASE FROM
TITLE OF INVENTION: A FUNGUS INFECTED ALGAE, ITS PURIFICATION
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knudde, Martens, Olsson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/633,768A
FILING DATE: 02-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9321301.5
FILING DATE: 15-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E.
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: DY007.001APC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714-760-0404
TELEFAX: 714-760-9502
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1088 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-633-768A-1

Query Match 3.6%; Score 86; DB 3; Length 1088;
Best Local Similarity 21.6%; Pred. No. 8.4; 138; Indels 84; Gaps 15;
Matches 72; Conservative 40; Mismatches 138; Indels 84; Gaps 15;

QY 106 VHGTPAV-IKTKRSILDETIMQTIKGPQVYQETDPIFLMGNEKMLTKKDAKOLE 164
DB 29 IHSAFPVHATRTKTNLN--VSMTALSDKOTAFAGSTD-----NDGI---DYTYD 76
QY 165 YAAK-QETPLS-----ISFDLDRNNTPLWSSRPHNPMYVLP 202
DB 77 YVGWGFSPLSNNTWFAAGSSTPGITDWTAFMNVNEDRIDNP--STVQHPVOVQVTSY 134
QY 203 HGKPNRSNTPSHARQFTNEFRAPRLKFOYSKVAKAEDLMGDSLMTGYTQOSH 262
DB 135 NNSSTVRFPNDGIRVTV---RGPIKQOLD-WITQELSECDPGM----- 178
QY 263 IFNGKNSRPVRVDYQPEIFLTQPVYSDLPWDGKVRMIGMCAVHSHSGESAKLSRWNR 322
DB 179 TETSEGLTFETKDSLVIYIIGNFRTVTRKSDGKVIN-----ENDEVGTAS----- 224
QY 323 YLMAGMEKMLTYVPRIMGRIFKEGSGSDPDNDILDIYGYGDVRFLYO---LENKSN 379
DB 225 ---SGNCRGLMEFYDRLYGNALIASVKNFRNDKAVKQEGFYGAGEVCKYODTYILERTGI 281
QY 380 SGT-----VRYN-----PRSGKALQLDLVYPL 402
DB 282 AMTNYNDNLNMQMDLRPHHDGALNPDYIPM 315

Search completed: November 30, 2001, 14:17:48
Job time: 225 sec

...

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:24:38 ; Search time 358.98 Seconds
(Without alignments) 341.870 Million cell updates/sec

Title: US-09-787-083-6

Perfect score: 2360
Sequence: 1 MKVSLSTLTSLTSCFALIA.....YNHEATSGVGLMDWML 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

Pending Patents AA Main: *
1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US080_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2360	100.0	442	21	US-09-787-083-6
2	2352	99.7	442	21	US-09-787-083-2
3	2348	99.5	442	21	US-09-787-083-8
4	2330	98.7	442	21	US-09-787-083-4
5	1810	76.7	346	19	US-09-540-236-3202
6	1785	75.6	375	24	US-60-128-476-3019
7	815	34.5	374	21	US-09-787-084-4
8	812	34.4	375	21	US-09-787-084-2
9	552	23.4	109	19	US-09-540-236-3197

	10	552	23.4	132	24	US-60-128-476-3016	Sequence 3016, Ap
	11	548.5	23.2	388	17	US-09-328-352-7532	Sequence 7532, Ap
	12	372.5	15.8	293	17	US-09-543-681A-4390	Sequence 4390, Ap
	13	371	15.7	297	18	US-09-489-039A-8021	Sequence 8021, Ap
	14	357	15.1	263	21	US-09-758-460-912	Sequence 912, Ap
	15	354	15.0	307	16	US-09-252-691-7841	Sequence 7841, Ap
	16	354	15.0	307	16	US-09-252-691-7841	Sequence 7841, Ap
	17	352.5	14.9	282	24	US-60-215-161-7051	Sequence 7051, Ap
	18	263	11.1	749	1	PCT-US01-08631-55161	Sequence 55161, A
	19	246.5	10.4	355	1	PCT-US98-06371-1342	Sequence 1342, Ap
	20	246.5	10.4	355	13	US-08-902-615A-608	Sequence 608, Ap
	21	239.5	10.1	356	11	PCT-US97-19575-90	Sequence 90, Ap
	22	239.5	10.1	356	11	US-08-759-739-472	Sequence 472, Ap
	23	239.5	10.1	356	13	US-08-993-002A-5205	Sequence 5205, Ap
	24	173.5	7.4	591	1	PCT-US01-08631-55155	Sequence 55155, A
	25	156.5	6.6	251	11	US-08-759-739-504	Sequence 504, Ap
	26	156.5	6.6	251	13	US-08-993-002A-5203	Sequence 5203, Ap
	27	156.5	6.6	253	11	US-08-759-739-381	Sequence 381, Ap
	28	156.5	6.6	253	13	US-08-993-002A-5204	Sequence 5204, Ap
	29	127.5	5.4	134	10	US-08-625-811-1275	Sequence 1275, Ap
	30	114	4.8	1686	17	US-09-355-160A-2	Sequence 2, Ap
	31	114	4.8	1686	17	US-09-355-160B-2	Sequence 2, Ap
	32	107	4.5	157	8	US-08-487-032A-837	Sequence 837, Ap
	33	107	4.5	157	9	US-08-561-469A-837	Sequence 837, Ap
	34	107	4.5	157	11	US-08-759-739-327	Sequence 327, Ap
	35	107	4.5	157	13	US-08-993-002A-5202	Sequence 5202, Ap
	36	104.5	4.4	537	21	US-09-738-626-4754	Sequence 4754, Ap
	37	104.5	4.4	546	20	US-09-603-124B-410	Sequence 410, Ap
	38	104.5	4.4	1030	24	US-60-324-109-30815	Sequence 30815, A
	39	104	4.4	798	16	PCT-US99-28332-5	Sequence 5, Ap
	40	104	4.4	798	16	US-09-203-453-5	Sequence 5, Ap
	41	104	4.4	798	21	US-09-712-032-5	Sequence 5, Ap
	42	104	4.4	798	23	US-09-948-947-86	Sequence 86, Ap
	43	104	4.4	1414	24	US-60-314-050-7662	Sequence 7662, Ap
	44	101.5	4.3	467	1	PCT-US01-04098A-1004	Sequence 1004, Ap
	45	101.5	4.3	467	14	US-09-062-009-6	Sequence 6, Ap

ALIGNMENTS

RESULT 1
US-09-787-083-6
; Sequence 6, Application US/09787083
; GENERAL INFORMATION:
; APPLICANT: Ruelie, Jean-Louis
; TITLE OF INVENTION: Novel Compounds
; FILE REFERENCE: BM45332
; CURRENT APPLICATION NUMBER: US/09/787,083
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: PCT/EP99/06781
; PRIOR FILING DATE: 1999-09-14
; PRIOR APPLICATION NUMBER: GB 9820002.5
; PRIOR FILING DATE: 1998-09-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-09-787-083-6

Query Match 100.0%; Score 2360; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 4.2e-235;
Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MKVSLSTLTSLTSCFALIAQAKAVNPVAFVDEVSNDLGDNELPTDVOSATQSA 60
Db 1 MKVSLSTLTSLTSCFALIAQAKAVNPVAFVDEVSNDLGDNELPTDVOSATQSA 60
OY 61 STDPAHPDEHPPELYTALNKMTLINCALNDIMLACYDPLVHGEPFAVITKTKRSI 120
|||||

Db 61 STDTPANPLDEHEBELYTTALENKTMLINCSALNODIMRLACYDTLVHGETPAVIKTKRSI 120
QY 121 RLDFTIMQTIKGRPOVYVOETDTPIFLMGNEKMLTKKDAKOLEYAKOFTPLSLSFDD 180
Db 121 RLDFTIMQTIKGRPOVYVOETDTPIFLMGNEKMLTKKDAKOLEYAKOFTPLSLSFDD 180
QY 181 RNNTPPLMSRPHNPMYVLPFIEMHGKPNRSPNTPSHAKOFTPNEFRAPBLKFOVSYKVA 240
Db 181 RNNTPPLMSRPHNPMYVLPFIEMHGKPNRSPNTPSHAKOFTPNEFRAPBLKFOVSYKVA 240
QY 241 ADDLMGTDSDLMEGYTQOSHMOIFNGKNSRPFVHDYOEIETLTOPVYSDLPMDGKVRMI 300
Db 241 ADDLMGTDSDLMEGYTQOSHMOIFNGKNSRPFVHDYOEIETLTOPVYSDLPMDGKVRMI 300
QY 301 GMAVAHNSGESAKLSRSNNRAYLMAEMKMLTVMPRIWGRIFKSGSGSQPDNDPDIID 360
Db 301 GMAVAHNSGESAKLSRSNNRAYLMAEMKMLTVMPRIWGRIFKSGSGSQPDNDPDIID 360
QY 361 YYGCGVRFYLOLEKNSNIGTVRYNPRSGKALQDLYVYPLGKISGYFOIFOGYGOSL 420
Db 361 YYGCGVRFYLOLEKNSNIGTVRYNPRSGKALQDLYVYPLGKISGYFOIFOGYGOSL 420
QY 421 IDYNHEATSEFGVGLMLNDMMGL 442
Db 421 IDYNHEATSEFGVGLMLNDMMGL 442

RESULT 2

US-09-787-083-2
; Sequence 2, Application us/09787083
; GENERAL INFORMATION:
; APPLICANT: Ruelle, Jean-Louis
; FILE OF INVENTION: Novel Compounds
; FILE REFERENCE: BM45332
; CURRENT APPLICATION NUMBER: US/09/787,083
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: PCT/EP99/06781
; PRIOR FILING DATE: 1999-09-14
; PRIOR APPLICATION NUMBER: GB 9820002.5
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-09-787-083-2

Query Match 99.7%; Score 2352; DB 21; Length 442;

Best Local Similarity 99.3%; Pred. No. 2,8e-234;

Matches 439; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSLSTLTLISLSCFAIILAIQAKAVPNPVAFVDEVRENDGONELPIDVQATQSA 60
Db 1 MKVSLSTLTLISLSCFAIILAIQAKAVPNPVAFVDEVRENDGONELPIDVQATQSA 60
QY 61 STDTPANPLDEHEBELYTTALENKTMLINCSALNODIMRLACYDTLVHGETPAVIKTKRSI 120
Db 61 STDTPANPLDEHEBELYTTALENKTMLINCSALNODIMRLACYDTLVHGETPAVIKTKRSI 120
QY 121 RLDFTIMQTIKGRPOVYVOETDTPIFLMGNEKMLTKKDAKOLEYAKOFTPLSLSFDD 180
Db 121 RLDFTIMQTIKGRPOVYVOETDTPIFLMGNEKMLTKKDAKOLEYAKOFTPLSLSFDD 180
QY 181 RNNTPPLMSRPHNPMYVLPFIEMHGKPNRSPNTPSHAKOFTPNEFRAPBLKFOVSYKVA 240
Db 181 RNNTPPLMSRPHNPMYVLPFIEMHGKPNRSPNTPSHAKOFTPNEFRAPBLKFOVSYKVA 240
QY 241 ADDLMGTDSDLMEGYTQOSHMOIFNGKNSRPFVHDYOEIETLTOPVYSDLPMDGKVRMI 300
Db 241 ADDLMGTDSDLMEGYTQOSHMOIFNGKNSRPFVHDYOEIETLTOPVYSDLPMDGKVRMI 300

QY 301 GMAVAHNSGESAKLSRSNNRAYLMAEMKMLTVMPRIWGRIFKSGSGSQPDNDPDIID 360
Db 301 GMAVAHNSGESAKLSRSNNRAYLMAEMKMLTVMPRIWGRIFKSGSGSQPDNDPDIID 360
QY 361 YYGCGVRFYLOLEKNSNIGTVRYNPRSGKALQDLYVYPLGKISGYFOIFOGYGOSL 420
Db 361 YYGCGVRFYLOLEKNSNIGTVRYNPRSGKALQDLYVYPLGKISGYFOIFOGYGOSL 420
QY 421 IDYNHEATSEFGVGLMLNDMMGL 442
Db 421 IDYNHEATSEFGVGLMLNDMMGL 442

RESULT 3

US-09-787-083-8
; Sequence 8, Application us/09787083
; GENERAL INFORMATION:
; APPLICANT: Ruelle, Jean-Louis
; FILE OF INVENTION: Novel Compounds
; FILE REFERENCE: BM45332
; CURRENT APPLICATION NUMBER: US/09/787,083
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: PCT/EP99/06781
; PRIOR FILING DATE: 1999-09-14
; PRIOR APPLICATION NUMBER: GB 9820002.5
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-09-787-083-8

Query Match 99.5%; Score 2348; DB 21; Length 442;

Best Local Similarity 99.3%; Pred. No. 7,3e-234;

Matches 439; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKVSLSTLTLISLSCFAIILAIQAKAVPNPVAFVDEVRENDGONELPIDVQATQSA 60
Db 1 MKVSLSTLTLISLSCFAIILAIQAKAVPNPVAFVDEVRENDGONELPIDVQATQSA 60
QY 61 STDTPANPLDEHEBELYTTALENKTMLINCSALNODIMRLACYDTLVHGETPAVIKTKRSI 120
Db 61 STDTPANPLDEHEBELYTTALENKTMLINCSALNODIMRLACYDTLVHGETPAVIKTKRSI 120
QY 121 RLDFTIMQTIKGRPOVYVOETDTPIFLMGNEKMLTKKDAKOLEYAKOFTPLSLSFDD 180
Db 121 RLDFTIMQTIKGRPOVYVOETDTPIFLMGNEKMLTKKDAKOLEYAKOFTPLSLSFDD 180
QY 181 RNNTPPLMSRPHNPMYVLPFIEMHGKPNRSPNTPSHAKOFTPNEFRAPBLKFOVSYKVA 240
Db 181 RNNTPPLMSRPHNPMYVLPFIEMHGKPNRSPNTPSHAKOFTPNEFRAPBLKFOVSYKVA 240
QY 241 ADDLMGTDSDLMEGYTQOSHMOIFNGKNSRPFVHDYOEIETLTOPVYSDLPMDGKVRMI 300
Db 241 ADDLMGTDSDLMEGYTQOSHMOIFNGKNSRPFVHDYOEIETLTOPVYSDLPMDGKVRMI 300
QY 301 GMAVAHNSGESAKLSRSNNRAYLMAEMKMLTVMPRIWGRIFKSGSGSQPDNDPDIID 360
Db 301 GMAVAHNSGESAKLSRSNNRAYLMAEMKMLTVMPRIWGRIFKSGSGSQPDNDPDIID 360
QY 361 YYGCGVRFYLOLEKNSNIGTVRYNPRSGKALQDLYVYPLGKISGYFOIFOGYGOSL 420
Db 361 YYGCGVRFYLOLEKNSNIGTVRYNPRSGKALQDLYVYPLGKISGYFOIFOGYGOSL 420
QY 421 IDYNHEATSEFGVGLMLNDMMGL 442
Db 421 IDYNHEATSEFGVGLMLNDMMGL 442

RESULT 4

```

US-09-787-083-4
: Sequence 4, Application US/09787083
:
: GENERAL INFORMATION:
:
: APPLICANT: Kuelie, Jean-Louis
:
: TITLE OF INVENTION: Novel Compounds
:
: FILE REFERENCE: BM45332
:
: CURRENT APPLICATION NUMBER: US/09/787,083
:
: CURRENT FILING DATE: 2001-03-14
:
: PRIOR APPLICATION NUMBER: PCT/EP99/06781
:
: PRIOR FILING DATE: 1999-09-14
:
: PRIOR APPLICATION NUMBER: GB 9820002.5
:
: PRIOR FILING DATE: 1998-09-14
:
: NUMBER OF SEQ ID NOS: 13
:
: SOFTWARE: FastSeq for Windows Version 4.0
:
: SEQ ID NO 4
:
: LENGTH: 442
:
: TYPE: PRT
:
: ORGANISM: Moraxella catarrhalis
:
: US-09-787-083-4

```

Query Match	98.7%;	Score 2330;	DB 21;	Length 442;
Best Local Similarity	98.9%;	Pred. NO. 5.4e-232;		
Matches 437; Conservative	2;	Mismatches 3;	Indels 0;	Gaps 0;

QY	1	MKVSLSTLSTLSLSCFPIIALAQAKAAPNPYAFVDEVRSENDJGQDNELPLDVQASQSA	60
Db	1	MKVSLSTLSTLSLSCFPIIALAQAKAAPNPYAFVDEVRSENDJGQDNELPLDVQASQSA	60
QY	61	STDTANPLDEHERELYTATALENKTMTLINGSALNODIMRLACYPDLVNGETPAVYIKTKRSI	120
Db	61	STDTANPLDEHERELYTATALENKTMTLINGSALNODIMRLACYPDLVNGETPAVYIKTKRSI	120
QY	121	RLDFTIQTIGKRGQVYVYQETPDRIFLIMGNEKGYLTCKDAKQLEYAKQTPPLSLSTDL	180
Db	121	RLDFTIQTIGKRGQVYVYQETPDRIFLIMGNEKGYLTCKDAKQLEYAKQTPPLSLSTDL	180
QY	181	RNNRPLMSSRRHNEMUYLPIRFMHKRPKRSPTPSHEARQETPNEFRAPELKFGVSYVKA	240
Db	181	RNNRPLMSSRRHNEMUYLPIRFMHKRPKRSPTPSHEARQETPNEFRAPELKFGVSYVKA	240
QY	241	AEDLMGTDSDLMFEGYTOQSHWQIFNGKNSRPFVNDYQREPLFLOPYSDLPMDGKYRMT	300
Db	241	AEDLMGTDSDLMFEGYTOQSHWQIFNGKNSRPFVNDYQREPLFLOPYSDLPMDGKYRMT	300
QY	301	GMGAVHNSNGESALTSRWNAUYLMAAMEKNLTVMPRIGRIFKEBSSGSGOPDNPDIID	360
Db	301	GMGAVHNSNGESALTSRWNAUYLMAAMEKNLTVMPRIGRIFKEBSSGSGOPDNPDIID	360
QY	361	YVYGADVAFVLEOLENKNSNISGTVVYNNRSGKALQILDYVYPLGKGISGYQIFQGYQSL	420
Db	361	YVYGADVAFVLEOLENKNSNISGTVVYNNRSGKALQILDYVYPLGKGISGYQIFQGYQSL	420
QY	421	IDYVHHEATSPGVGLMDNMGL	442
Db	421	IDYVHHEATSPGVGLMDNMGL	442

```

RESULT      5
US-09-540-236-3202
: Sequence 3202, Application US/09540236
: GENERAL INFORMATION:
: APPLICANT: Gary L. Breton et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
: TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 2709.2005-001
: CURRENT APPLICATION NUMBER: US/09/540.236
: CURRENT FILING DATE: 2000-04-04
: NUMBER OF SEQ ID NOS: 3840
: SEQ ID NO 3202
: LENGTH: 346
: TYPE: PRT
: ORGANISM: M.catarrhalis

```

US-09-540-236-3202

Query Match	76.7%;	Score 1810;	DB 19;	Length 346;
Best Local Similarity	99.4%;	Pred. No. 3.2e-178;		
Matches 339;	Conservative 2;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	MKVSLSTLLTSLISCFALITAOQKAVNPAFDEVSENDLGDONDELPIDVOSAQOSA	60
Db	3	MKVSLSTLLTSLISCFALITAOQKAVNPAFDEVSENDLGDONDELPIDVOSAQOSA	62
QY	61	STDTPANPLDEHPELYTTALENKTMLNCSALNODIMRLACYDPLVHGEPYAVIKTKRSI	120
Db	63	STDTPANPLDEHPELYTTALENKTMLNCSALNODIMRLACYDPLVHGEPYAVIKTKRSI	122
QY	121	RLDERTIMOTIGKRPVVYQOETTPDPIFLMGNSKGMILTKDKAQLEYAKOFTPLSTSLD	180
Db	123	RLDERTIMOTIGKRPVYIQOETTPDPIFLMGNSKGMILTKDKAQLEYAKOFTPLSTSLD	182
QY	181	RNNTPPLMSSRRHPNPAUYLPRTIMHGKPNPSPTPSHEARQFTPNBEFRAPBLKFOYSUYVKA	240
Db	183	RNNTPPLMSSRRHPNPAUYLPRTIMHGKPNPSPTPSHEARQFTPNBEFRAPBLKFOYSUYVKA	242
QY	241	AEDLWGTSIDLMFGYTOQSHNQIFENGKNSRPFRRVHDQPELTELQPAYUSDLPMDGKYRMI	300
Db	243	AEDLWGTSIDLMFGYTOQSHNQIFENGKNSRPFRRVHDQPELTELQPAYUSDLPMDGKYRMI	302
QY	301	GMGAVHSHNGESAKLSRSWNPAYLTMAHGEMKNLTVMPRTWG	341
Db	303	GMGAVHSHNGESAKLSRSWNPAYLTMAHGEMKNLTVMPRTWG	343

```

RESULT 6
US-60-128-476-3019
; Sequence 3019, Application US/60128476
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA
; TITLE OF INVENTION: CATARRHALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.139
; CURRENT APPLICATION NUMBER: US/60/128.476
; CURRENT FILING DATE: 1998-04-09
; NUMBER OF SEQ ID NOS: 5002
; SEQ ID NO 3019
; LENGTH: 375
; TYPE: PRF
; ORGANISM: Moraxella catarrhalis
US-60-128-476-3019

```

Query Match	75.6%;	Score 1785;	DB 24;	Length 375;
Best Local Similarity	98.0%;	Pred. No. 1.5e-175;		
Matches 336;	Conservative 3;	Mismatches 4;	Indels 0;	Gaps 0;

Qy	1	MKSLSTLTLSLSEFALIAQAAKAVNPAPAFDEVRESENDJGODNHELPIDVQASQSA	60
Db	3	MKVSJSTLTLSLSEFALIAQAAAPNPAPAFDEVRESENDJGODNHELPIDVQASQSA	62
Qy	61	STDTANPLDEHEBPELTYTALENKTMILNCSALNDIMKLACTDYLVGGETPAVTKTRSI	120
Db	63	STDNANPLDEHEBPELTYTALENKTMILNCSALNDIMKLACTDYLVGGETPAVTKTRSI	122
Qy	121	RLDETWTOTIGKKPOVYVYQETTTDPLFLMGNNKGMILTKDAKQLEYAAKQTPPLSSTDLD	180
Db	123	RLDETWTOTIGKKPOVYVYQETTTDPLFLMGNNKGMILTKDAKQLEYAAKQTPPLSSTDLD	182
Qy	181	RNNPLMSSRRHPNRYVYLPFLFMCKGKPNSSPMTPSHEAKQTPNBEFRADELKFOVSYVKA	240
Db	183	RNNPLMSSRRHPNRYVYLPFLFMCKGKPNSSPMTPSHEAKQTPNBEFRADELKFOVSYVKA	242
Qy	241	AEDLMGTDSDLMFGYTOOSHMOIFNGKNRPFVHVHYQPELITLQPYVSDLPMDGKVRMI	300
Db	243	AEDLMGTDSDLMFGYTOOSHMOIFNGKNRPFVHVHYQPELITLQPYVSDLPMDGKVRMI	302

```
QY      301 GMGAVHSHNGESAKLSRSWNRAYLMAGMEWKNTVMPRIWGRI 343  
        |||||  
Db      303 GMGAVHSHNGESAKLSRSWNRAYLMAGMEWKNTVMPHTLGGV 345
```

RESULT 7

```

US-09-787-084-4
: Sequence 4, Application US/09787084
:
: GENERAL INFORMATION:
: APPLICANT: Ruelle, Jean-Louis
: TITLE OF INVENTION: Polynucleotides and Polypeptides BASB03
: TITLE OF INVENTION: from Neisseria Meningitidis and Their Uses
: FILE REFERENCE: BM45331
: CURRENT APPLICATION NUMBER: US/09/787, 084
: CURRENT FILING DATE: 2001-03-14
: PRIOR APPLICATION NUMBER: PCT/EP99/06718
: PRIOR FILING DATE: 1999-09-09
: PRIOR APPLICATION NUMBER: GB 9820003.3
: PRIOR FILING DATE: 1998-09-14
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 4
:
: LENGTH: 374
:
: TYPE: PRT
:
: ORGANISM: Neisseria meningitidis
: US-09-787-084-4

```

Query Match	34.58;	Score 815;	DB 21;	Length 374;
-------------	--------	------------	--------	-------------

QY	87	INCALNODIRRLACCYDTLVHGETPAVI-----KTKRSIRLDETIMOTI-KGRPOVYOE	140
Db	25	LQCALLDINVTIRLACCYDRIFRAQDLPSSAGOBGEOSKAVLMTETVSSLSLKGAEVIVYEK	84
QY	141	TTDPDIFLMNGEKMGLTKKAKOLEYAAKOTPLSLSPDLDRNN-TPLMSSRPNNPVLP	199
Db	85	GGDAL-----PADSAGETADITYTPLSLMTDLDKNDRLGLGVRHNNPYMLP	131
QY	200	IFMHGKPKRSNPNTSHEAR-QFTTNERABELKFQVSVKYKAAEDLMTGSDLMFGYTQO	258
Db	132	LMYNNSPIYAVGSPTRGTITVOEKGGQOKRAETKLOVSEKSIIEDELFTKTRADLMFGYTQR	191
QY	259	SHMOIFN-GKMSRPFRRHNDOPRELFLOPYUSDLPMDGKVRMIGMAGVHNSNESAKLSR	317
Db	192	SDMOITVMOGRKSAFPRNTDKPRTEFLQPVKADLPFGSRRLMGLAGVHNSNOSPRESR	251
QY	318	SMNRAYVLAMGEMENLTVMPRIWGRIFKEGSGSQPDNPDLIDVGYGADVRFYOLENKS	377
Db	252	SMNRITVYAMAGHEMKRLTVIPRVVWRADQ--SGDK-NONPRIDADYMGVDVAKLOYLRLNDRQ	309
QY	378	NISGTVRYNPRSIGALQLDLYVVRPLGKISGYEQIIFQGYGOSLIDYVNEHETSFGVGLMLN	437
Db	310	NVYSVLRYNPRXTGYATEAATPFIKGLKGVVYGFHGYGSLIDYVNHKNGIGIGLMFN	369
QY	438	DMHGL 442	
Db	370	DLDGI 374	

RESULT 8

US-09-787-084-2
Sequence 2, Application US/09787084
GENERAL INFORMATION:
APPLICANT: Ruelle, Jean-Louis
TITLE OF INVENTION: Polynucleotides and Polypeptides BASB033
TITLE OF INVENTION: from Neisseria Meningitidis and Their Uses
FILE REFERENCE: BM45331
CURRENT APPLICATION NUMBER: US/09/787, 084
CURRENT FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: PCT/EP99/06718

```

: PRIOR FILING DATE: 1999-03-09
: PRIOR APPLICATION NUMBER: GB 9820003.3
: PRIOR FILING DATE: 1998-03-14
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: FastSeq for Windows Version 4.0.

```

ORGANISM: *Neisseria meningitidis*
US-09-787-084-2

QY	87	INCSALNODIRRLACYPDTLVHGEPPAVI-----KTKRSIRLDETIMOTI-KGRPOVYOE	140
Db	26	LOCALDLNDVTRLCYURITRAQDLPSSAGOBGQESKAVLMLTETVRSLSLKGEAVIYEK	85
QY	141	TTDPDIFLMGNEKMLTKKDAKOLEYAAKQETPLSLSPDLDRNN-TPLMSSRPNNHVPYLR	199
Db	86	GGDAL-----PADSAGETADITYTPLSLMWDLDKNDLRGLGVREHNNPMLP	132
QY	200	IFMHGKPNRSNPPTSHEAR-QFTPRNEFRADELKFQVYSVKKAABDLMTGSDMLFGYTQO	258
Db	133	LMYNNSPNYAPSSPTRGTTQOEKGQOKRAETKLOVYSFKSIEDLEKTRADMLFGYTQR	192
QY	259	SHMOIEN-GKNSRPFRRNDVORELELORPYYSDLPWMOGKVRMIQMGVHNHNSNEASKLR	317
Db	193	SDMOITVNOGKRSAPFRNRTDYKPELETLQPVKADLPFGGRULMIGAGVUHOSNGQSPRES	252
QY	318	SMNRAUYLMAGMEKMLTVMPRLIMGRIFKEGSGSQPDNPDLIDYUGGVDFRFLTOLENS	377
Db	253	SMNRIYAMAGMECKLIVIRVWVRARDQ--SGDR-NONPRIDADMGVDVKLQYRLNDRQ	310
QY	378	NISGTAVRNPBSKAGALODLVYVPLGKIGISGYFOIFQYQOSLIDVNHNEATSFVGGLMLN	437
Db	311	NVYSLVRNPNPTGALIEAAVTFPIKGLKGVUVGFHGYGESLIDYVNHKONGIGIGIMFN	370
QY	438	DMWGL 442	
Db	371	DLGGI 375	

RESULT 9

```

US-09-540-236-3197
; Sequence 3197, Application US/09540236
; GENERAL INFORMATION:
; APPLICANT: Gary L. Bretton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709. 2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3197
; LENGTH: 109
; TYPE: PRN
; ORGANISM: M.cattarrhalis
; US-09-540-236-3197

```

Query Match	23.4%	Score 553	DB 19	Length 109
Best Local Similarity	100.0%	Pred. NO.	1.7e-48	
Matches 102	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY 341 GRKESSGSPDDNPILLDYGYGVDYRLFQLEKNSISGTVRYNRRSGKGLQLDYVY 400
|||||
Db 8 GRKESSGSPDDNPILLDYGYGVDYRLFQLEKNSISGTVRYNRRSGKGLQLDYVY 67
|||||
QY 401 PLKGISGTFDIFQGYQSLLDYNHEHTSPGVGLMLDNMGL 442
|||||

Db 68 PLKGISGYFOIFGCGSLIDYHNEATSFVGGLMDMMGL 109

RESULT 10
US-60-128-476-3016

; Sequence 3016, Application US/60128476

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA

; FILE REFERENCE: 107196.139

; CURRENT APPLICATION NUMBER: US/60/128,476

; CURRENT FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 5002

; SEQ ID NO 3016

; LENGTH: 132

; TYPE: PR

; ORGANISM: Moraxella catarrhalis

US-60-128-476-3016

Query Match

23.4%; Score 552; DB 24; Length 132;

Best Local Similarity 100.0%; Pred. No. 2.3e-48;

Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 341 GRIFKESGSDPDNDPILDYGYGVFLYOLENKSINISGTVYRNRSGKALQLDIVY 400

Db 31 GRIFKESGSDPDNDPILDYGYGVFLYOLENKSINISGTVYRNRSGKALQLDIVY 90

Db 401 PLKGISGYFOIFGCGSLIDYHNEATSFVGGLMDMMGL 442

Db 91 PLKGISGYFOIFGCGSLIDYHNEATSFVGGLMDMMGL 132

RESULT 11

US-09-328-352-7532

; Sequence 7532, Application US/09328352

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; FILE REFERENCE: 57C99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 7532

; LENGTH: 388

; TYPE: PR

; ORGANISM: Acinetobacter baumannii

US-09-328-352-7532

Query Match

23.2%; Score 548.5; DB 17; Length 388;

Best Local Similarity 34.4%; Pred. No. 3e-47;

Matches 126; Conservative 70; Mismatches 129; Indels 41; Gaps 10;

Db 89 CSALINDIMRLACYDVLHGETPAVIKT---KRSIRLDETWTQIKGPOVYQETTPD 144

Db 48 CVALANSADRLACYDAVFKPSALPVQAAVPEPVKKIDRPVQEFTEKVV--DKVSN 105

Db 145 IFLMGNKMKILTKDAQOLEAKQFTPLSLSP-----LDRNNTPLMSSPHNPKVLP 199

Db 106 IKVYGR-----AFTLEPTSLDQWELSEKSLGVWNIKAYQPYLLP 149

Db 200 IFMHGKPNRPNTPSHEAROFTPNEFAPELKFOVSKVKAEDLMGTDSDLPFGYTOQS 259

Db 150 VFTMSDKNEFPSPNPNNTYEAONLKSSTESKQISLKTAMENIFGNGDLMVGYTOQS 209

Db 260 HMOIFNGKNSRPFVHDYQPE---IFLTQVYSDLPMDGKVRMTGKAVHSHNSGSAKLS 316

Db 210 RMQTFNAEESRPFRETNYEPEASLMEFTN--YELLGLDG--RLDGVTLLHNSGNSRDPDS 265

Db 317 RSMNRATILMAGMEKNTLVPRIGRIFKESGSDPDNDPILDYGYGVFLYOLENK 376

Db 266 RSMNRVFNVLGKGNRPLMRPWIREE---SKDDNPNPMEDYIGRGDLATATYKK-Q 321

Db 377 SNISGTVYRNRSG---KCALQLDIVYPLKGISGYFOIFGCGSLIDYHNEATSFVG 433

Db 322 NDFSLMRHSLKGGDDSHGAVQFDMAPFISGLKRGHFLGFGYGESLIDYHNRATYAGIG 381

Db 434 LMLNDM 439

Db 382 VSLMNV 387

RESULT 12

US-09-543-681A-4390

; Sequence 4390, Application US/09543681A

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRA

; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/09/543,681A

; CURRENT FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 60/128,706

; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 8344

; SEQ ID NO 4390

; LENGTH: 293

; TYPE: PR

; ORGANISM: Proteus mirabilis

US-09-543-681A-4390

Query Match

15.8%; Score 372.5; DB 19; Length 293;

Best Local Similarity 39.6%; Pred. No. 3.4e-29;

Matches 90; Conservative 35; Mismatches 87; Indels 15; Gaps 6;

Db 215 HEAQFTPNERRAPLEKFQVSVKKAEDLMGTSDLMFTGTOOSHQOIFNGKSRPRPV 274

Db 81 HKAK-----DEVKFOISLAEPLWNGIAGENSVAASTORSMMOLSKRESAPPR 132

Db 275 HDYQPEIFLQPV-YSDLPMDGKVRMTGMAVHSHNSGESAKLSRSMNRATILMAGMENL 333

Db 133 TNYEPOFLGNTDYQFAGW--TLREITGTFHNSNGAETSTSMNRVYRFAQGNF 190

Db 334 TVMPRIWGRIFKESGSDPDNDPILDYGYGVFLYOLENKSINISGTVYRNRSGKGA 393

Db 191 QLDLKPWYR-FNE--SAKNDNDPINDRYMGYRRLKAGYRL-GESEFTLIGRYNMSSGYGA 246

Db 394 QLDIVYPLKGISGYFOIFGCGSLIDYHNEATSFVGGLMDMM 440

Db 247 AELGWSYPTIKHVRPYTQLFSGYGESMIDYNFROTRLIGVGLMDML 293

RESULT 13

US-09-489-039A-8021

; Sequence 8021, Application US/09489039A

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 8021

; LENGTH: 297

; TYPE: PR

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-8021

Inis Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:55:42 ; Search time 36.87 Seconds
(Without alignments)
269.771 Million cell updates/sec

Title: US-09-787-083-8

Perfect score: 442

Sequence: 1 MKVSLSTLTLSPCFALIA.....YVHEATSPGVGLMDNMCL 442

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 212252 seqs, 22503292 residues

Word size : 0

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCBUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	1.6	16	1	US-07-942-245-98 Sequence 98, Appl
2	7	1.6	552	1	US-08-231-7298-6 Sequence 6, Appl
3	7	1.6	900	4	US-08-890-865A-4 Sequence 4, Appl
4	7	1.6	934	1	US-08-215-805A-80 Sequence 80, Appl
5	7	1.6	1026	2	US-08-342-003-6 Sequence 6, Appl
6	7	1.6	1026	2	US-08-322-760A-6 Sequence 6, Appl
7	6	1.4	10	3	US-08-974-775-30 Sequence 30, Appl
8	6	1.4	11	3	US-08-974-775-29 Sequence 29, Appl
9	6	1.4	12	3	US-08-974-775-28 Sequence 28, Appl
10	6	1.4	13	3	US-08-974-775-8 Sequence 8, Appl
11	6	1.4	13	3	US-08-974-775-27 Sequence 27, Appl
12	6	1.4	14	3	US-08-974-775-9 Sequence 9, Appl
13	6	1.4	14	3	US-08-974-775-10 Sequence 10, Appl
14	6	1.4	14	3	US-08-974-775-32 Sequence 32, Appl
15	6	1.4	15	3	US-08-974-775-5 Sequence 5, Appl
16	6	1.4	16	3	US-08-974-775-4 Sequence 4, Appl
17	6	1.4	20	1	US-08-430-273-1 Sequence 1, Appl
18	6	1.4	21	1	US-07-918-953-13 Sequence 13, Appl
19	6	1.4	21	1	US-07-918-953-15 Sequence 15, Appl
20	6	1.4	21	1	US-08-212-696-1 Sequence 1, Appl
21	6	1.4	21	1	US-08-158-245-1 Sequence 1, Appl
22	6	1.4	21	1	US-08-081-661-13 Sequence 13, Appl
23	6	1.4	21	1	US-08-081-661-15 Sequence 15, Appl
24	6	1.4	21	1	US-08-233-617-1 Sequence 1, Appl
25	6	1.4	21	1	US-08-160-376A-1 Sequence 1, Appl
26	6	1.4	21	1	US-08-304-070-1 Sequence 1, Appl
27	6	1.4	21	1	US-08-285-661-1 Sequence 1, Appl

28	6	1.4	21	1	US-08-301-838-1 Sequence 1, Appl
29	6	1.4	21	1	US-08-389-487-4 Sequence 4, Appl
30	6	1.4	21	1	US-08-389-487-12 Sequence 12, Appl
31	6	1.4	21	1	US-08-507-124-2 Sequence 2, Appl
32	6	1.4	21	1	US-08-507-124-3 Sequence 3, Appl
33	6	1.4	21	1	US-08-342-931-1 Sequence 1, Appl
34	6	1.4	21	1	US-08-400-256-1 Sequence 1, Appl
35	6	1.4	21	2	US-08-508-664-9 Sequence 9, Appl
36	6	1.4	21	2	US-08-353-476-85 Sequence 85, Appl
37	6	1.4	21	2	US-08-353-476-87 Sequence 87, Appl
38	6	1.4	21	2	US-08-484-219-1 Sequence 1, Appl
39	6	1.4	21	2	US-08-979-587-1 Sequence 1, Appl
40	6	1.4	21	2	US-08-992-676-1 Sequence 1, Appl
41	6	1.4	21	2	US-08-982-676-6 Sequence 6, Appl
42	6	1.4	21	2	US-09-134-836-1 Sequence 1, Appl
43	6	1.4	21	2	US-09-134-836-7 Sequence 7, Appl
44	6	1.4	21	3	US-08-967-867-2 Sequence 2, Appl
45	6	1.4	21	4	US-08-932-082-1 Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-07-942-245-98
Sequence 98, Application US/07942245
Patent No. 5639641
GENERAL INFORMATION:
APPLICANT: PEDERSEN, Jan T.
APPLICANT: SEARLE, Stephen M.J.
APPLICANT: REES, Anthony R.
APPLICANT: ROGUSKA, Michael A.
APPLICANT: GUILD, Braydon C.
TITLE OF INVENTION: SURFACE RESIDUE VENEERING OF RODENT
TITLE OF INVENTION: ANTIBODIES
NUMBER OF SEQUENCES: 522
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zimm, Macpreak & Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: HP 9000/700 Workstation
OPERATING SYSTEM: UNIX
SOFTWARE: In house
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/942,245
FILING DATE: 09-SEP-1992
CLASSIFICATION: 530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 98:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-942-245-98

Query Match 1.6%; Score 7; DB 1; Length 16;
Best local Similarity 100.0%; Pred. No. 5.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
QY 387 PRSGGA 393
|||||
Db 4 PRSGGA 10

RESULT 2
US-08-231-729B-6
; Sequence 6, Application US/08231729B
; Patent No. 5618722
; GENERAL INFORMATION:
; APPLICANT: ZENNO, Shuhei
; APPLICANT: SHIRAIISHI, Shinji
; APPLICANT: INOUE, Satoshi
; APPLICANT: SAIGO, Kaoru
; TITLE OF INVENTION: FIRELY LUCIFERASE GENE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEYDIG, VOIT & MAYER
; STREET: 700 Thirteenth Street, N.W., Suite 300
; City: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,729B
; FILING DATE: 20-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 119050/1993
; FILING DATE: 21-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rose, Herbert C.
; REGISTRATION NUMBER: 29846
; REFERENCE/DOCKET NUMBER: 60130/No. 5618722aka
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-737-6770
; TELEFAX: 202-737-6776
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 552 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-231-729B-6

Query Match 1.6%; Score 7; DB 1; Length 552;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 LSPFLDR 181
|111111|
Db 184 LSPFLDR 190

RESULT 3
US-08-890-865A-4
; Sequence 4, Application US/08890865A
; Patent No. 6307019
; GENERAL INFORMATION:
; APPLICANT: Constantini, Franklin
; APPLICANT: Zeng, Li
; TITLE OF INVENTION: AXIN GENE AND USES THEREOF
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; City: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10036
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/890,865A
; FILING DATE: 10-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/54249
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)278-0400
; TELEFAX: (212)391-0526
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 900 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-890-865A-4

Query Match 1.6%; Score 7; DB 4; Length 900;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 403 GKIGISG 409
|1111111|
Db 268 GKIGISG 274

RESULT 4
US-08-215-805A-80
; Sequence 80, Application US/08215805A
; Patent No. 5559008
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; TITLE OF INVENTION: LEUKOTOXIN GENE FROM PASTEURRELLA
; TITLE OF INVENTION: SUIS
; NUMBER OF SEQUENCES: 84
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; City: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/215,805A
; FILING DATE: 22-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Timlan, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/61 (D-1329A)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1636
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 934 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein

ORIGINAL SOURCE:
ORGANISM: Pasteurella suis
STRAIN: 5943
IMMEDIATE SOURCE:
LIBRARY: P. suis DNA in Bacteriophage lambda-dash
CLONE: (Lambda)yfc3-37
US-08-215-805A-80

Query Match 1.6%; Score 7; DB 1; Length 934;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 238 VKAEDL 244
|||||
Db 62 VKAEDL 68

RESULT 5
US-08-542-003-6
Sequence 6, Application US/08542003
Patent No. 5864013
GENERAL INFORMATION:
APPLICANT: Goldberg, Edward B.
TITLE OF INVENTION: MATERIALS FOR THE PRODUCTION OF
TITLE OF INVENTION: NANOMETER STRUCTURES AND USE THEREOF
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennle and Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/542.003
FILING DATE: 13-OCT-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 8471-0005-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: 212-869-8864
TELEX: 66441 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1026 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Bacteriophage T4
IMMEDIATE SOURCE:
CLONE: p37 amino acid
US-08-542-003-6

Query Match 1.6%; Score 7; DB 2; Length 1026;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 128 OTIKGP 134
|||||
Db 861 OTIKGP 867

RESULT 6
US-08-322-760A-6
Sequence 6, Application US/08322760A
Patent No. 5877279
GENERAL INFORMATION:
APPLICANT: Goldberg, Edward B.
TITLE OF INVENTION: MATERIALS FOR THE PRODUCTION OF
TITLE OF INVENTION: NANOMETER STRUCTURES AND USE THEREOF
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennle and Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/322.760A
FILING DATE: 13-OCT-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 8471-0003-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: 212-869-8864
TELEX: 66441 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1026 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Bacteriophage T4
IMMEDIATE SOURCE:
CLONE: p37 amino acid
US-08-322-760A-6

Query Match 1.6%; Score 7; DB 2; Length 1026;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 128 OTIKGP 134
|||||
Db 861 OTIKGP 867

RESULT 7
US-08-974-775-30
Sequence 30, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
APPLICANT: Toback, F. Gary
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
STREET: Plaza Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775
FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-30
```

```
Query Match 1.4%; Score 6; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 424 NHEATS 429
      |||||
DB 5 NHEATS 10
```

```
RESULT 8
US-08-974-775-29
Sequence 29, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
APPLICANT: Toback, F. Gary
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
STREET: Plaza Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775
FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
```

```
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-29
```

```
Query Match 1.4%; Score 6; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 424 NHEATS 429
      |||||
DB 5 NHEATS 10
```

```
RESULT 9
US-08-974-775-28
Sequence 28, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
APPLICANT: Toback, F. Gary
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
STREET: Plaza Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775
FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-28
```

```
Query Match 1.4%; Score 6; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 424 NHEATS 429
      |||||
DB 5 NHEATS 10
```

```
RESULT 10
US-08-974-775-8
Sequence 8, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
```

```
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; STREET: Plaza Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-775-8

Query Match          1.4%; Score 6; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
      |||||
      8 NHEATS 13

Db

RESULT 11
US-08-974-775-27
; Sequence 27, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; STREET: Plaza Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
```

```
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-775-27

Query Match          1.4%; Score 6; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
      |||||
      5 NHEATS 10

Db

RESULT 12
US-08-974-775-9
; Sequence 9, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; STREET: Plaza Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-775-9

Query Match          1.4%; Score 6; DB 3; Length 14;
```

Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 424 NHEATS 429
|||||
Db 8 NHEATS 13

RESULT 13
US-08-974-775-10
; Sequence 10, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; STREET: Plaza Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ. ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-974-775-10

Query Match 1.4%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 424 NHEATS 429
|||||
Db 8 NHEATS 13

RESULT 14
US-08-974-775-32
; Sequence 32, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE

; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; STREET: Plaza Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ. ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE: Modified-site
; NAME/KEY: 14
; LOCATION: 14
; OTHER INFORMATION: /product="Ser or Tyr"
US-08-974-775-32

Query Match 1.4%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 424 NHEATS 429
|||||
Db 8 NHEATS 13

RESULT 15
US-08-974-775-5
; Sequence 5, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; STREET: Plaza Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:

NAME: Martin, Alice O.
 REGISTRATION NUMBER: 35,601
 REFERENCE/DOCKET NUMBER: 7814/27
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-321-4200
 TELEFAX: 312-321-4299
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 15 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-974-775-5

Query Match 1.4%; Score 6; DB 3; Length 15;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 424 NHEATS 429
 Db 8 NHEATS 13

Search completed: May 9, 2002, 12:55:43
 Job time: 553 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 13:03:24 ; Search time 403.82 Seconds
(Without alignments)
384,464 Million cell updates/sec

Title: US-09-787-083-8
Perfect score: 442
Sequence: 1 MKVSLSTLTLSTLPCFALIA.....YVHEATSPGVGLMDNMCL 442

Scoring table:
Gapex 60.0 , Gapext 60.0

Searched: 3516493 seqs, 351254056 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3516493

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/PCTUS.COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06.COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07.COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US080.COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US081.COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082.COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083.COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084.COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085.COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086.COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087.COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088.COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089.COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090.COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091.COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092.COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093.COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094.COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095.COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096.COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097.COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US098.COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US099.COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US60.COMB.pep.*
25: /cgn2_6/ptodata/2/paa/US100.MERGED.COMB.pep.*
26: /cgn2_6/ptodata/2/paa/US101.MERGED.COMB.pep.*
27: /cgn2_6/ptodata/2/paa/US102.MERGED.COMB.pep.*
28: /cgn2_6/ptodata/2/paa/US103.MERGED.COMB.pep.*
29: /cgn2_6/ptodata/2/paa/US104.MERGED.COMB.pep.*
30: /cgn2_6/ptodata/2/paa/US105.MERGED.COMB.pep.*
31: /cgn2_6/ptodata/2/paa/US106.MERGED.COMB.pep.*
32: /cgn2_6/ptodata/2/paa/US107.MERGED.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	442	100.0	442	21	US-09-787-083-8

2	327	74.0	442	21	US-09-787-083-2	Sequence 2, Appl
3	316	71.5	442	21 <th>US-09-787-083-6</th> <th>Sequence 6, Appl</th>	US-09-787-083-6	Sequence 6, Appl
4	289	65.4	442	21 <th>US-09-787-083-4</th> <th>Sequence 4, Appl</th>	US-09-787-083-4	Sequence 4, Appl
5	215	48.6	346	19 <th>US-09-540-236-3202</th> <th>Sequence 3202, Ap</th>	US-09-540-236-3202	Sequence 3202, Ap
6	211	47.7	375	24 <th>US-60-128-476-3019</th> <th>Sequence 3019, Ap</th>	US-60-128-476-3019	Sequence 3019, Ap
7	102	23.1	109	9 <th>US-09-540-236-3197</th> <th>Sequence 3197, Ap</th>	US-09-540-236-3197	Sequence 3197, Ap
8	102	23.1	132	24 <th>US-60-128-476-3016</th> <th>Sequence 3016, Ap</th>	US-60-128-476-3016	Sequence 3016, Ap
9	9	2.0	255	1 <th>PCT-US01-08631-48393</th> <th>Sequence 48393, A</th>	PCT-US01-08631-48393	Sequence 48393, A
10	9	2.0	374	21 <th>US-09-787-084-4</th> <th>Sequence 4, Appl</th>	US-09-787-084-4	Sequence 4, Appl
11	9	2.0	375	21 <th>US-09-787-084-2</th> <th>Sequence 2, Appl</th>	US-09-787-084-2	Sequence 2, Appl
12	8	1.8	227	25 <th>US-60-361-742-1681</th> <th>Sequence 1681, Ap</th>	US-60-361-742-1681	Sequence 1681, Ap
13	8	1.8	297	18 <th>US-09-489-039A-8021</th> <th>Sequence 8021, Ap</th>	US-09-489-039A-8021	Sequence 8021, Ap
14	7	1.6	50	15 <th>US-09-161-939-8</th> <th>Sequence 8, Appl</th>	US-09-161-939-8	Sequence 8, Appl
15	7	1.6	50	15 <th>US-09-161-939A-8</th> <th>Sequence 8, Appl</th>	US-09-161-939A-8	Sequence 8, Appl
16	7	1.6	52	21 <th>US-09-760-498-674</th> <th>Sequence 674, App</th>	US-09-760-498-674	Sequence 674, App
17	7	1.6	58	24 <th>US-60-170-429-1266</th> <th>Sequence 1266, Ap</th>	US-60-170-429-1266	Sequence 1266, Ap
18	7	1.6	58	24 <th>US-60-182-568-1001</th> <th>Sequence 1001, Ap</th>	US-60-182-568-1001	Sequence 1001, Ap
19	7	1.6	65	19 <th>US-09-595-329A-1580</th> <th>Sequence 1580, Ap</th>	US-09-595-329A-1580	Sequence 1580, Ap
20	7	1.6	70	18 <th>US-09-417-507-23275</th> <th>Sequence 23275, A</th>	US-09-417-507-23275	Sequence 23275, A
21	7	1.6	73	24 <th>US-60-160-203-4820</th> <th>Sequence 4820, Ap</th>	US-60-160-203-4820	Sequence 4820, Ap
22	7	1.6	73	24 <th>US-60-162-247-4285</th> <th>Sequence 4285, Ap</th>	US-60-162-247-4285	Sequence 4285, Ap
23	7	1.6	73	24 <th>US-60-169-840-6621</th> <th>Sequence 6621, Ap</th>	US-60-169-840-6621	Sequence 6621, Ap
24	7	1.6	75	1 <th>PCT-US01-01340-1013</th> <th>Sequence 1013, Ap</th>	PCT-US01-01340-1013	Sequence 1013, Ap
25	7	1.6	75	21 <th>US-09-764-869-1013</th> <th>Sequence 1013, Ap</th>	US-09-764-869-1013	Sequence 1013, Ap
26	7	1.6	75	26 <th>US-10-091-504-1013</th> <th>Sequence 1013, Ap</th>	US-10-091-504-1013	Sequence 1013, Ap
27	7	1.6	85	24 <th>US-60-188-162-4426</th> <th>Sequence 4426, Ap</th>	US-60-188-162-4426	Sequence 4426, Ap
28	7	1.6	91	27 <th>US-09-708-427-41947</th> <th>Sequence 41947, A</th>	US-09-708-427-41947	Sequence 41947, A
29	7	1.6	99	24 <th>US-60-187-387-1177</th> <th>Sequence 1177, Ap</th>	US-60-187-387-1177	Sequence 1177, Ap
30	7	1.6	107	27 <th>US-09-594-595B-888</th> <th>Sequence 888, App</th>	US-09-594-595B-888	Sequence 888, App
31	7	1.6	116	1 <th>PCT-US01-01384-72</th> <th>Sequence 72, Appl</th>	PCT-US01-01384-72	Sequence 72, Appl
32	7	1.6	116	1 <th>PCT-US01-01384-99</th> <th>Sequence 99, Appl</th>	PCT-US01-01384-99	Sequence 99, Appl
33	7	1.6	116	1 <th>PCT-US01-01384-100</th> <th>Sequence 100, Appl</th>	PCT-US01-01384-100	Sequence 100, Appl
34	7	1.6	116	1 <th>PCT-US01-11988-1993</th> <th>Sequence 1993, App</th>	PCT-US01-11988-1993	Sequence 1993, App
35	7	1.6	116	1 <th>PCT-US01-11988-1954</th> <th>Sequence 1954, Ap</th>	PCT-US01-11988-1954	Sequence 1954, Ap
36	7	1.6	116	1 <th>PCT-US01-11988-1955</th> <th>Sequence 1955, Ap</th>	PCT-US01-11988-1955	Sequence 1955, Ap
37	7	1.6	116	22 <th>US-09-833-245-1953</th> <th>Sequence 1953, Ap</th>	US-09-833-245-1953	Sequence 1953, Ap
38	7	1.6	116	22 <th>US-09-833-245-1954</th> <th>Sequence 1954, Ap</th>	US-09-833-245-1954	Sequence 1954, Ap
39	7	1.6	116	22 <th>US-09-833-245-1955</th> <th>Sequence 1955, Ap</th>	US-09-833-245-1955	Sequence 1955, Ap
40	7	1.6	133	1 <th>PCT-US01-08631-57915</th> <th>Sequence 57915, A</th>	PCT-US01-08631-57915	Sequence 57915, A
41	7	1.6	134	1 <th>PCT-US01-14827-13005</th> <th>Sequence 13005, A</th>	PCT-US01-14827-13005	Sequence 13005, A
42	7	1.6	156	18 <th>US-09-489-039A-12780</th> <th>Sequence 12780, A</th>	US-09-489-039A-12780	Sequence 12780, A
43	7	1.6	158	15 <th>US-09-138-133-42</th> <th>Sequence 42, Appl</th>	US-09-138-133-42	Sequence 42, Appl
44	7	1.6	158	21 <th>US-09-789-919-42</th> <th>Sequence 42, Appl</th>	US-09-789-919-42	Sequence 42, Appl
45	7	1.6	160	18 <th>US-09-489-039A-12847</th> <th>Sequence 12847, A</th>	US-09-489-039A-12847	Sequence 12847, A

ALIGNMENTS

RESULT 1
US-09-787-083-8
Sequence 8, Application US/09787083
GENERAL INFORMATION:
APPLICANT: Ruelle, Jean-Louis
TITLE OF INVENTION: Novel Compounds
FILE REFERENCE: BM45332
CURRENT APPLICATION NUMBER: US/09/787, 083
CURRENT FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: PCT/EP99/06781
PRIOR FILING DATE: 1999-09-14
PRIOR APPLICATION NUMBER: GB 9820002.5
PRIOR FILING DATE: 1998-09-14
NUMBER OF SEQ ID NOS: 13
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 442
TYPE: PRT
ORGANISM: Moraxella catarrhalis
US-09-787-083-8

Query Match 100.0%; Score 442; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 0;

```
Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MKVSLSTLTLILPCFAILAIQQAQAVNPVAFVDEVRENDLGODNELPIDVOSATQSA 60
Db 1 MKVSLSTLTLILPCFAILAIQQAQAVNPVAFVDEVRENDLGODNELPIDVOSATQSA 60
Qy 61 STDTANPLDEHEBELYTTALENTMLINCALNODIMRLACYDTLVHGETPAVYIKTKRSI 120
Db 61 STDTANPLDEHEBELYTTALENTMLINCALNODIMRLACYDTLVHGETPAVYIKTKRSI 120
Qy 121 RLDEETIMOTIKGRPOVYVOETTPFILMGNEKMLTKKDAKOLEYAKOFTPLSLFDLD 180
Db 121 RLDEETIMOTIKGRPOVYVOETTPFILMGNEKMLTKKDAKOLEYAKOFTPLSLFDLD 180
Qy 181 RNNTPLMSSRPHPMPYVLPFLMHGKPNRSPNTPSHEAKOFTPNFRAPELKFQSVKVA 240
Db 181 RNNTPLMSSRPHPMPYVLPFLMHGKPNRSPNTPSHEAKOFTPNFRAPELKFQSVKVA 240
Qy 241 AEDLMGTDSLWMFGYTOQSHMOIFNGKNSRPFVHNDYQPEIFLTOPYSDLPMDGKVRMI 300
Db 241 AEDLMGTDSLWMFGYTOQSHMOIFNGKNSRPFVHNDYQPEIFLTOPYSDLPMDGKVRMI 300
Qy 301 GMGAVHHSNGESAALRSRNRAIYLMAGMEKMLTVMPRIKGRIFKEGSGSQPDNDPILD 360
Db 301 GMGAVHHSNGESAALRSRNRAIYLMAGMEKMLTVMPRIKGRIFKEGSGSQPDNDPILD 360
Qy 361 YYGVDVRFPLYOLENKSNIQVRYNPRSGGALQLDYVYPLKGISGYFOIFQYGOQL 420
Db 361 YYGVDVRFPLYOLENKSNIQVRYNPRSGGALQLDYVYPLKGISGYFOIFQYGOQL 420
Qy 421 IDYNHEATSGVGLMDNMGL 442
Db 421 IDYNHEATSGVGLMDNMGL 442

RESULT 2
US-09-787-083-2
; Sequence 2, Application US/09787083
; GENERAL INFORMATION:
; APPLICANT: Ruelle, Jean-Louis
; TITLE OF INVENTION: Novel Compounds
; FILE REFERENCE: BM45332
; CURRENT APPLICATION NUMBER: US/09/787,083
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: PCT/EP99/06781
; PRIOR FILING DATE: 1999-09-14
; PRIOR APPLICATION NUMBER: GB 9820002.5
; PRIOR FILING DATE: 1998-09-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 442
; TYPE: PRF
; ORGANISM: Moraxella catarrhalis
US-09-787-083-2
```

```
Query Match 74.0%; Score 327; DB 21; Length 442;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 442; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 15 CFAILAIQQAQAVNPVAFVDEVRENDLGODNELPIDVOSATQSASTDTANPLDEHEPE 74
Db 15 CFAILAIQQAQAVNPVAFVDEVRENDLGODNELPIDVOSATQSASTDTANPLDEHEPE 74
Qy 75 LYTALLENKMLINCALNODIMRLACYDTLVHGETPAVYIKTKRSIRLDEETIMOTIKGR 134
Db 75 LYTALLENKMLINCALNODIMRLACYDTLVHGETPAVYIKTKRSIRLDEETIMOTIKGR 134
Qy 135 QVYVOETTPFILMGNEKMLTKKDAKOLEYAKOFTPLSLFDLDNRNNTPLMSSRPHP 194
Db 135 QVYVOETTPFILMGNEKMLTKKDAKOLEYAKOFTPLSLFDLDNRNNTPLMSSRPHP 194
```

```
Qy 195 MYVLPFLMHGKPNRSPNTPSHEAKOFTPNFRAPELKFQSVKVAEDLMGTDSLWMFG 254
Db 195 MYVLPFLMHGKPNRSPNTPSHEAKOFTPNFRAPELKFQSVKVAEDLMGTDSLWMFG 254
Qy 255 YTOQSHMOIFNGKNSRPFVHNDYQPEIFLTOPYSDLPMDGKVRMIGMGAVHHSNGESA 314
Db 255 YTOQSHMOIFNGKNSRPFVHNDYQPEIFLTOPYSDLPMDGKVRMIGMGAVHHSNGESA 314
Qy 315 LSRSNRAIYLMAGMEKMLTVMPRIKGRIFKEGSGSQPDNDPILDYGYGVDFPLYOLE 374
Db 315 LSRSNRAIYLMAGMEKMLTVMPRIKGRIFKEGSGSQPDNDPILDYGYGVDFPLYOLE 374
Qy 375 NKSNIQVRYNPRSGGALQLDYVYPLKGISGYFOIFQYGOQLIDYNHEATSGVGL 434
Db 375 NKSNIQVRYNPRSGGALQLDYVYPLKGISGYFOIFQYGOQLIDYNHEATSGVGL 434
Qy 435 MLNDMMGL 442
Db 435 MLNDMMGL 442
```

```
RESULT 3
US-09-787-083-6
; Sequence 6, Application US/09787083
; GENERAL INFORMATION:
; APPLICANT: Ruelle, Jean-Louis
; TITLE OF INVENTION: Novel Compounds
; FILE REFERENCE: BM45332
; CURRENT APPLICATION NUMBER: US/09/787,083
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: PCT/EP99/06781
; PRIOR FILING DATE: 1999-09-14
; PRIOR APPLICATION NUMBER: GB 9820002.5
; PRIOR FILING DATE: 1998-09-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 442
; TYPE: PRF
; ORGANISM: Moraxella catarrhalis
US-09-787-083-6
```

```
Query Match 71.5%; Score 316; DB 21; Length 442;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 26 AVNPVAFVDEVRENDLGODNELPIDVOSATQSASTDTANPLDEHEPELYTTALENTKM 85
Db 26 AVNPVAFVDEVRENDLGODNELPIDVOSATQSASTDTANPLDEHEPELYTTALENTKM 85
Qy 86 LINCALNODIMRLACYDTLVHGETPAVYIKTKRSIRLDEETIMOTIKGRPOVYVOETTPDI 145
Db 86 LINCALNODIMRLACYDTLVHGETPAVYIKTKRSIRLDEETIMOTIKGRPOVYVOETTPDI 145
Qy 146 FLMGNEKMLTKKDAKOLEYAKOFTPLSLFDLDNRNNTPLMSSRPHPMPYVLPFLMHGK 205
Db 146 FLMGNEKMLTKKDAKOLEYAKOFTPLSLFDLDNRNNTPLMSSRPHPMPYVLPFLMHGK 205
Qy 206 PNRSPNTPSHEAKOFTPNFRAPELKFQSVKVAEDLMGTDSLWMFGYTOQSHMOIQIN 265
Db 206 PNRSPNTPSHEAKOFTPNFRAPELKFQSVKVAEDLMGTDSLWMFGYTOQSHMOIQIN 265
Qy 266 GKNSRPFVHNDYQPEIFLTOPYSDLPMDGKVRMIGMGAVHHSNGESAALRSRNRAIYLM 325
Db 266 GKNSRPFVHNDYQPEIFLTOPYSDLPMDGKVRMIGMGAVHHSNGESAALRSRNRAIYLM 325
Qy 326 AGMEKMLTVMPRIKGRIFKEGSGSQPDNDPILDYGYGVDFPLYOLENKSNIQVRY 385
Db 326 AGMEKMLTVMPRIKGRIFKEGSGSQPDNDPILDYGYGVDFPLYOLENKSNIQVRY 385
Qy 386 NPRSGGALQLDYVYPLKGISGYFOIFQYGOQLIDYNHEATSGVGLMDNMGL 442
Db 386 NPRSGGALQLDYVYPLKGISGYFOIFQYGOQLIDYNHEATSGVGLMDNMGL 442
```

Db 386 NPSRGKALQLDYVYPLGKGISGYFQIFQYGGQSLIDYNHEATSGVGLMLNDMMGL 442

RESULT 4
US-09-787-083-4
: Sequence 4, Application US/09787083
: GENERAL INFORMATION:
: APPLICANT: Ruelle, Jean-Louis
: TITLE OF INVENTION: Novel Compounds
: FILE REFERENCE: BM45332
: CURRENT APPLICATION NUMBER: US/09/787,083
: CURRENT FILING DATE: 2001-03-14
: PRIOR APPLICATION NUMBER: PCT/EP99/06781
: PRIOR FILING DATE: 1999-09-14
: PRIOR APPLICATION NUMBER: GB 9820002.5
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 4
: LENGTH: 442
: TYPE: PRT
: ORGANISM: Moraxella catarrhalis
US-09-787-083-4

Query Match 65.4%; Score 289; DB 21; Length 442;
Best Local Similarity 99.7%; Pred. No. 2,3e-291;
Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 53 VQSATQASSTDPANPLDEHEPELYTTALENTMLINCSALNODIMRLACYDTLVHGETPA 112
DB 53 VQSATQASSTDPANPLDEHEPELYTTALENTMLINCSALNODIMRLACYDTLVHGETPA 112
QY 113 VIKTKRSIRLDETTMOTIKGRPOVYVOTETDPIFLMGNEKMLTKKAKOLEYAKOFTP 172
DB 113 VIKTKRSIRLDETTMOTIKGRPOVYVOTETDPIFLMGNEKMLTKKAKOLEYAKOFTP 172
QY 173 LSLFSDLDNRNTPRLSSRPNRPVYLPFLFMGKRPKRSPTNSHEKAKOTPEFRAPELKEF 232
DB 173 LSLFSDLDNRNTPRLSSRPNRPVYLPFLFMGKRPKRSPTNSHEKAKOTPEFRAPELKEF 232
QY 233 QVSKVKAARDLMDGTDSDLMFGYTOQSHMOJFNGKNSRPFRVNDOPETFLTOPVYSDLP 292
DB 233 QVSKVKAARDLMDGTDSDLMFGYTOQSHMOJFNGKNSRPFRVNDOPETFLTOPVYSDLP 292
QY 293 WDGKVRMIGMGAHVHNSGESAKLSRSNNRAYLTMAGMEKMLTVMPRIKRIEKGSGSQP 352
DB 293 WDGKVRMIGMGAHVHNSGESAKLSRSNNRAYLTMAGMEKMLTVMPRIKRIEKGSGSQP 352
QY 353 DDNDPDIIDYGYGVRFYLYOLEKNSNISGYTRYNPRSGKALQLDYVYPLGKGISGYFOI 412
DB 353 DDNDPDIIDYGYGVRFYLYOLEKNSNISGYTRYNPRSGKALQLDYVYPLGKGISGYFOI 412
QY 413 FQGYGQSLIDYNHEATSGVGLMLNDMMGL 442
DB 413 FQGYGQSLIDYNHEATSGVGLMLNDMMGL 442

RESULT 5
US-09-540-236-3202
: Sequence 3202, Application US/09540236
: GENERAL INFORMATION:
: APPLICANT: Gary L. Breton et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
: FILE REFERENCE: 2709, 2005-001
: CURRENT APPLICATION NUMBER: US/09/540,236
: CURRENT FILING DATE: 2000-04-04
: NUMBER OF SEQ ID NOS: 3840
: SEQ ID NO 3202
: LENGTH: 346
: TYPE: PRT
: ORGANISM: M. catarrhalis

US-09-540-236-3202

Query Match 48.6%; Score 215; DB 19; Length 346;
Best Local Similarity 99.7%; Pred. No. 2,1e-214;
Matches 315; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 26 AVPNPVAFVDEVRSNDLGDONELPIDVQATQASSTDPANPLDEHEPELYTTALENTM 85
DB 26 AVPNPVAFVDEVRSNDLGDONELPIDVQATQASSTDPANPLDEHEPELYTTALENTM 85
QY 86 LINCASALNODIMRLACYDTLVHGETPAVYIKTKRSIRLDETTMOTIKGRPOVYVOTETDPI 145
DB 86 LINCASALNODIMRLACYDTLVHGETPAVYIKTKRSIRLDETTMOTIKGRPOVYVOTETDPI 145
QY 146 FLMGNEKMLTKKAKOLEYAKOFTPLSLFSDLDNRNTPRLSSRPNRPVYLPFLFMHGK 205
DB 146 FLMGNEKMLTKKAKOLEYAKOFTPLSLFSDLDNRNTPRLSSRPNRPVYLPFLFMHGK 205
QY 206 PNRSPNTPSHAKOFTPEFRAPELKFOVSVKAAEDLMGTDSDLMFGYTOQSHMOJFN 265
DB 206 PNRSPNTPSHAKOFTPEFRAPELKFOVSVKAAEDLMGTDSDLMFGYTOQSHMOJFN 265
QY 268 GKNSRPFRVNDYOPETFLTOPVYSDLPWDGKVRMIGMGAHVHNSGESAKLSRSNNRAYLM 327
DB 268 GKNSRPFRVNDYOPETFLTOPVYSDLPWDGKVRMIGMGAHVHNSGESAKLSRSNNRAYLM 327
QY 326 AGMEKMLTVMPRIWG 341
DB 326 AGMEKMLTVMPRIWG 341

RESULT 6
US-60-128-476-3019
: Sequence 3019, Application US/60128476
: GENERAL INFORMATION:
: APPLICANT: Gary L. Breton et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196,139
: CURRENT APPLICATION NUMBER: US/60/128,476
: CURRENT FILING DATE: 1999-04-09
: NUMBER OF SEQ ID NOS: 5002
: SEQ ID NO 3019
: LENGTH: 375
: TYPE: PRT
: ORGANISM: Moraxella catarrhalis
US-60-128-476-3019

Query Match 47.7%; Score 211; DB 24; Length 375;
Best Local Similarity 99.7%; Pred. No. 3,3e-210;
Matches 311; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 26 AVPNPVAFVDEVRSNDLGDONELPIDVQATQASSTDPANPLDEHEPELYTTALENTM 85
DB 26 AVPNPVAFVDEVRSNDLGDONELPIDVQATQASSTDPANPLDEHEPELYTTALENTM 85
QY 86 LINCASALNODIMRLACYDTLVHGETPAVYIKTKRSIRLDETTMOTIKGRPOVYVOTETDPI 145
DB 86 LINCASALNODIMRLACYDTLVHGETPAVYIKTKRSIRLDETTMOTIKGRPOVYVOTETDPI 145
QY 146 FLMGNEKMLTKKAKOLEYAKOFTPLSLFSDLDNRNTPRLSSRPNRPVYLPFLFMHGK 205
DB 146 FLMGNEKMLTKKAKOLEYAKOFTPLSLFSDLDNRNTPRLSSRPNRPVYLPFLFMHGK 205
QY 206 PNRSPNTPSHAKOFTPEFRAPELKFOVSVKAAEDLMGTDSDLMFGYTOQSHMOJFN 265
DB 206 PNRSPNTPSHAKOFTPEFRAPELKFOVSVKAAEDLMGTDSDLMFGYTOQSHMOJFN 265
QY 268 GKNSRPFRVNDYOPETFLTOPVYSDLPWDGKVRMIGMGAHVHNSGESAKLSRSNNRAYLM 325
DB 268 GKNSRPFRVNDYOPETFLTOPVYSDLPWDGKVRMIGMGAHVHNSGESAKLSRSNNRAYLM 327

OY 326 AGMEMKNTVMP 337
Db 328 AGMEMKNTVMP 339

RESULT 7
US-09-540-236-3197

; Sequence 3197, Application US/09540236
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3197
; LENGTH: 109
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-3197

Query Match 23.1%; Score 102; DB 19; Length 109;
Best Local Similarity 100.0%; Pred. No. 3.3e-97;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 341 GRIFKESGSDPDNPILYGGVDFLYOLENKSISGTVARNRSGKALQLDYVY 400
Db 8 GRIFKESGSDPDNPILYGGVDFLYOLENKSISGTVARNRSGKALQLDYVY 67
OY 401 PLKGISGYFOIFQGYGOSLIDYNHEATSFVGIMLNDMGL 442
Db 68 PLKGISGYFOIFQGYGOSLIDYNHEATSFVGIMLNDMGL 109

RESULT 8
US-60-128-476-3016

; Sequence 3016, Application US/60128476
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.139
; CURRENT APPLICATION NUMBER: US/60/128,476
; CURRENT FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 5002
; SEQ ID NO 3016
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-60-128-476-3016

Query Match 23.1%; Score 102; DB 24; Length 132;
Best Local Similarity 100.0%; Pred. No. 3.9e-97;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 341 GRIFKESGSDPDNPILYGGVDFLYOLENKSISGTVARNRSGKALQLDYVY 400
Db 31 GRIFKESGSDPDNPILYGGVDFLYOLENKSISGTVARNRSGKALQLDYVY 90
OY 401 PLKGISGYFOIFQGYGOSLIDYNHEATSFVGIMLNDMGL 442
Db 91 PLKGISGYFOIFQGYGOSLIDYNHEATSFVGIMLNDMGL 132

RESULT 9
PCT-US01-08631-48393
; Sequence 48393, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc

; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 48393

; LENGTH: 255
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (135)..(152)
; OTHER INFORMATION: ArgE / dape / ACY1 / CPE2 / yacs family proteins domain
; OTHER INFORMATION: Identified by eMATRIX, accession number BL00738B, p-value=1.96
; NAME/KEY: DOMAIN
; LOCATION: (2)..(125)
; OTHER INFORMATION: Peptidase family M20/M25/M40 domain identified by Pfam,
; NAME/KEY: misc.feature
; LOCATION: (1)..(255)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-08631-48393

Query Match 2.0%; Score 9; DB 1; Length 255;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 PEIFLTQPV 287
Db 204 PEIFLTQPV 212

RESULT 10
US-09-787-084-4

; Sequence 4, Application US/09787084
; GENERAL INFORMATION:
; APPLICANT: Ruelie, Jean-Louis
; TITLE OF INVENTION: Polynucleotides and Polypeptides BASB033
; TITLE OF INVENTION: from Neisseria Meningitidis and Their Uses
; FILE REFERENCE: BM45331
; CURRENT APPLICATION NUMBER: US/09/787,084
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: PCT/EP99/06718
; PRIOR FILING DATE: 1999-09-09
; PRIOR APPLICATION NUMBER: GB 9820003.3
; PRIOR FILING DATE: 1998-09-14
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-787-084-4

Query Match 2.0%; Score 9; DB 21; Length 374;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 PEIFLTQPV 287
Db 213 PEIFLTQPV 221

RESULT 11
US-09-787-084-2

```
; Sequence 2, Application US/09787084
; GENERAL INFORMATION:
; APPLICANT: Rueille, Jean-Louis
; TITLE OF INVENTION: Polynucleotides and Polypeptides BASB033
; TITLE OF INVENTION: from Neisseria Meningitidis and Their Uses
; FILE REFERENCE: BM45331
; CURRENT APPLICATION NUMBER: US/09/787,084
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: PCT/EP99/06718
; PRIOR FILING DATE: 1999-09-09
; PRIOR APPLICATION NUMBER: GB 9820003.3
; PRIOR FILING DATE: 1998-09-14
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-787-084-2
```

```
Query Match          2.0%; Score 9; DB 21; Length 375;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 279 PEIFLTQPV 287
    |||||
DB 214 PEIFLTQPV 222
```

```
RESULT 12
US-60-361-742-1681
; Sequence 1681, Application US/60361742
; GENERAL INFORMATION:
; APPLICANT: Fidelity Systems, Inc.
; APPLICANT: Slesarev, Alexei I.
; APPLICANT: Mezhevaya, Katja V.
; APPLICANT: Polushin, Nikolai N.
; APPLICANT: Shcherbina, Olga V.
; APPLICANT: Shakhova, Vera V.
; APPLICANT: Malykh, Andrei G.
; APPLICANT: Kozavkin, Sergei A.
; TITLE OF INVENTION: Fidelity Systems, Inc.
; FILE REFERENCE: New
; CURRENT APPLICATION NUMBER: US/60/361,742
; CURRENT FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 1692
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1681
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Methanopyrus kandleri
US-60-361-742-1681
```

```
Query Match          1.8%; Score 8; DB 25; Length 227;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 105 LVHGETPA 112
    |||||
DB 106 LVHGETPA 113
```

```
RESULT 13
US-09-489-039A-8021
; Sequence 8021, Application US/09489039A
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
```

```
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8021
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8021
```

```
Query Match          1.8%; Score 8; DB 18; Length 297;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 431 GVGIMLND 438
    |||||
DB 288 GVGIMLND 295
```

```
RESULT 14
US-09-161-939-8
; Sequence 8, Application US/09161939
; GENERAL INFORMATION:
; APPLICANT: Shinkets, Richard A.
; TITLE OF INVENTION: Genes and Proteins Predictive and Therapeutic for
; FILE REFERENCE: 15966-527 SHR Genes and Proteins
; CURRENT APPLICATION NUMBER: US/09/161,939
; CURRENT FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-161-939-8
```

```
Query Match          1.6%; Score 7; DB 15; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 114 IKTKRSI 120
    |||||
DB 22 IKTKRSI 28
```

```
RESULT 15
US-09-161-939A-8
; Sequence 8, Application US/09161939A
; GENERAL INFORMATION:
; APPLICANT: Shinkets, Richard A.
; TITLE OF INVENTION: Genes and Proteins Predictive and Therapeutic for
; FILE REFERENCE: 15966-527
; CURRENT APPLICATION NUMBER: US/09/161,939A
; CURRENT FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-161-939A-8
```

```
Query Match          1.6%; Score 7; DB 15; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 114 IKTKRSI 120
    |||||
```

Db 22 IKTksi 28

Search completed: May 9, 2002, 13:03:25
Job time: 615 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:56:33 ; Search time 44.37 Seconds
(without alignments)
758.627 Million cell updates/sec

Title: US-09-787-083-6

Sequence: 442
1 MKVSLSTLTLSTLSCFAILA.....YNHEATSPGVGMLNDMMGL 442

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: PIR-68:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	2.0	382	2	E81195 phospholipase A1,
2	9	2.0	409	2	H81831 probable phospholip
3	8	1.8	286	2	B36971 outer membrane pho
4	8	1.8	299	2	T47989 RAV-like protein -
5	8	1.8	1396	2	S36851 L-shaped tail fibre
6	7	1.6	100	2	S61050 probable membrane
7	7	1.6	116	2	E86676 hypothetical prote
8	7	1.6	122	2	E75377 hypothetical prote
9	7	1.6	158	2	E86498 pts IIA protein (I
10	7	1.6	197	2	E72124 tubulin alpha-1 ch
11	7	1.6	219	2	T19438 hypothetical prote
12	7	1.6	237	2	C84976 tRNA (guanine-N1-)
13	7	1.6	240	2	A61544 tubulin alpha chai
14	7	1.6	264	2	A44959 coat protein - pot
15	7	1.6	267	2	A60366 coat protein - pot
16	7	1.6	267	2	S14001 genome polypotein
17	7	1.6	267	2	S26630 capsid protein - p
18	7	1.6	267	2	JC1527 coat protein - pot
19	7	1.6	267	2	S13239 tail fiber protein
20	7	1.6	270	2	S73734 abc transport ATP
21	7	1.6	273	2	C72338 transaminase B hom
22	7	1.6	287	2	A72421 oligopeptide ABC t
23	7	1.6	288	2	T36237 probable phosphotr
24	7	1.6	306	2	T27985 hypothetical prote
25	7	1.6	310	2	T02200 probable RAV-like
26	7	1.6	313	2	JT0960 polypotein - pota
27	7	1.6	321	2	H82240 conserved hypothet
28	7	1.6	324	2	T23876 hypothetical prote

30	7	1.6	326	2	A43939 proteinase inhibit
31	7	1.6	338	2	F69437 hypothetical prote
32	7	1.6	341	2	T47653 pectate lyase-like
33	7	1.6	342	2	G64411 hypothetical prote
34	7	1.6	342	2	C64394 hypothetical prote
35	7	1.6	347	1	DNHUN2 NADH dehydrogenase
36	7	1.6	347	2	T11051 NADH dehydrogenase
37	7	1.6	347	2	B59153 NADH dehydrogenase
38	7	1.6	349	2	T43920 yfuc protein [impo
39	7	1.6	370	2	G83219 probable dihydrol
40	7	1.6	379	2	S13556 genome polypotein
41	7	1.6	381	2	G82132 ribonuclease D VC1
42	7	1.6	382	2	S13237 tail fiber protein
43	7	1.6	384	2	S70638 tubulin alpha chai
44	7	1.6	384	2	S70641 tubulin alpha chai
45	7	1.6	384	2	S70640 tubulin alpha chai

ALIGNMENTS

RESULT 1

E81195 phospholipase A1, probable NMB0464 [imported] - Neisseria meningitidis (strain MC58 s

C:Species: Neisseria meningitidis

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C:Accession: E81195

R:Rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B. ri, H.; Qin, H.; Yamatehyan, J.; Gall, J.; Scariato, V.; Masignani, V.; Piazza, M. Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappunli, R.; A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755

A:Accession: E81195

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-382 <TEP>

A:Cross-references: GB:AE002403; GB:AE002098; NID:g7225688; PIDN:AAFA40901.1; PID:g7222

A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NMB0464

Query Match 2.0%; Score 9; DB 2; Length 382;

Best Local Similarity 100.0%; Pred. No. 0.46;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 279 PERFLTOPY 287

Db 221 PERFLTOPY 229

RESULT 2

H81831 probable phospholipase NMA2021 [imported] - Neisseria meningitidis (strain Z2491 serog

C:Species: Neisseria meningitidis

C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C:Accession: H81831

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491

A:Reference number: A81775; MUID:2022556

A:Accession: H81831

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-409 <PAR>

A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CA885240.1; PID:g738

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMA2021

Query Match 2.0%; Score 9; DB 2; Length 409;
 Best Local Similarity 100.0%; Pred. No. 0.49;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIEFLTPV 287
 |||||
 DB 248 PEIEFLTPV 256

RESULT 3
 B36971

Outer membrane phospholipase A (EC 3.1.1.-) precursor - *Klebsiella pneumoniae*
 C:Species: *Klebsiella pneumoniae*
 C>Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 18-Jun-1999
 C:Accession: B36971; S40129
 R:Brok, R.G.P.M.; Brinkman, E.; van Boxtel, R.; Bekkers, A.C.A.P.A.; Verheij, H.M.; Tomm
 J. Bacteriol 176, 861-870, 1994
 A>Title: Molecular characterization of enterobacterial pldA genes encoding outer membran
 A:Reference number: A36971; MUID:94131966
 A:Accession: B36971
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-286 <PRO>
 A:Cross-references: EMBL:X76901; NID:g436880; PID:CAA54223.1; PID:g436881
 A>Note: authors translated the codon AAG for residue 112 as Arg
 C:Genetics:
 A:Gene: pldA
 C:Superfamily: bacterial phospholipase A1
 C:Keywords: carboxylic ester hydrolase

Query Match 1.8%; Score 8; DB 2; Length 286;
 Best Local Similarity 100.0%; Pred. No. 4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 431 GVGGLMD 438
 |||||
 DB 277 GVGGLMD 284

RESULT 4
 T47989

RAV-like protein - *Arabidopsis thaliana*
 N:Alternate names: protein F21F14.140
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T47989
 R:Cholme, N.; Robert, C.; Brottier, P.; Wincker, P.; Catolico, L.; Artiguenave, F.; Sa
 submitted to the Protein Sequence Database, February 2000
 A:Reference number: Z24481
 A:Accession: T47989
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-299 <CHO>
 A:Cross-references: EMBL:AL138642
 A:Experimental source: cultivar Columbia; BAC clone F21F14
 C:Genetics:
 A:Map position: 3
 A>Note: F21F14.140

Query Match 1.8%; Score 8; DB 2; Length 299;
 Best Local Similarity 100.0%; Pred. No. 4.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 SLSPDDR 181
 |||||
 DB 292 SLSPDDR 299

RESULT 5
 S36851

L-shaped tail fiber protein - phage T5
 N:Alternate names: ltf protein
 C:Species: phage T5
 C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 17-Nov-2000
 C:Accession: S65934; S01984; S36851
 R:Kallman, A.V.; Kulshin, V.E.; Shlyupnikov, M.G.; Keenzenko, V.N.; Kryukov, V.M.
 FEBS Lett. 366, 46-48, 1995
 A>Title: The nucleotide sequence of the bacteriophage T5 ltf gene.
 A:Reference number: S65934; MUID:95309401
 A:Accession: S65934
 A:Molecule type: DNA
 A:Residues: 11396 <KAT>
 A:Cross-references: EMBL:X69460; NID:g15415; PID:CAA49220.1; PID:g15416
 R:Kallman, A.V.; Kryukov, V.M.; Bayev, A.A.
 Nucleic Acids Res. 16, 6230, 1988
 A>Title: The nucleotide sequence of bacteriophage T5 DNA at the region between early
 A:Reference number: S01982; MUID:88289370
 A:Accession: S01984
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 934-985, 'A', 987-1396 <KA2>
 A:Cross-references: EMBL:X07559
 C:Genetics:
 A:Gene: ltf
 C:Keywords: late protein; tail fiber

Query Match 1.8%; Score 8; DB 2; Length 1396;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 QSATQAS 61
 |||||
 DB 87 QSATQAS 94

RESULT 6
 S61050

Probable membrane protein YDL163w - yeast (*Saccharomyces cerevisiae*)
 N:Alternate names: hypothetical protein D1505
 C:Species: *Saccharomyces cerevisiae*
 C>Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 20-Jun-2000
 C:Accession: S61050; S67715
 R:Poehl, T.M.
 submitted to the EMBL Data Library, November 1995
 A:Reference number: S61010
 A:Accession: S61050
 A:Molecule type: DNA
 A:Residues: 1-100 <POH>
 A:Cross-references: EMBL:Z67750; NID:g1061256; PID:g1061276
 R:Poehl, T.M.
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S67708
 A:Accession: S67715
 A:Molecule type: DNA
 A:Residues: 1-100 <POW>
 A:Cross-references: EMBL:Z74212; NID:g1431258; PID:g1431260; MIPS:YDL163w
 A:Experimental source: strain S288C
 C:Genetics:
 A:Map position: 4L
 C:Superfamily: Saccharomyces probable membrane protein YDL163w
 C:Keywords: transmembrane protein
 F:3-19/Domain: transmembrane #status predicted <TM>

Query Match 1.6%; Score 7; DB 2; Length 100;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LTLSTLS 14
 |||||
 DB 26 LTLSTLS 32

RESULT 7
F86676
hypothetical protein ptca [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001
C:Accession: F86676
R:Boletín, A.; Mincker, P.; Manger, S.; Jallón, O.; Malarme, K.; Weissensbach, J.; Ehrlich
Genome Res. In press, 2001
A:Title: The complete genome sequence of the lactic acid bacterium.
A:Reference number: A86625
A:Accession: F86676
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-116 <STO>
A:Cross-references: GB:AE005176; NID:912723289; PIDN:AAK04512.1; GSPDB:GN00146
C:Genetics:
A:Gene: ptca

Query Match 1.6%; Score 7; DB 2; Length 116;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 A10QAKA 26
|||||
DB 34 A10QAKA 40

RESULT 8
E75377
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: E75377
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: E75377
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-122 <WHI>
A:Cross-references: GB:AE002002; GB:AE000513; NID:96459345; PIDN:AAF11150.1; PID:9645935
C:Genetics:
A:Experimental source: strain R1
A:Gene: DR1580
A:Map position: 1

Query Match 1.6%; Score 7; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 EHEPELY 76
|||||
DB 115 EHEPELY 121

RESULT 9
E86498
pts IIA protein [imported] - Chlamydomonas reinhardtii (strain J138)
C:Species: Chlamydomonas reinhardtii
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: E86498
R:Shirai, M.; Hirakawa, H.; Kinoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of Chlamydomonas reinhardtii J138.
A:Reference number: A86491; MUID:20330349
A:Accession: E86498

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-158 <STO>
A:Cross-references: GB:BA000008; NID:96978434; PIDN:BA98271.1; GSPDB:GN00142
C:Genetics:
A:Experimental source: strain J138
A:Gene: ptsN1
C:Superfamily: phosphotransferase system enzyme II; phosphotransferase system mannito

Query Match 1.6%; Score 7; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 LSTLTL 11
|||||
DB 126 LSTLTL 132

RESULT 10
E72124
pts IIA protein - Chlamydomonas reinhardtii (strain CWL029)
C:Species: Chlamydomonas reinhardtii
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
C:Accession: E72124
R:Kallman, S.; Mitchell, W.; Marathe, R.; Lamm, C.; Fan, J.; Olinger, L.; Grimwood,
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydomonas reinhardtii and C. trachomatis.
A:Reference number: A72000; MUID:99206606
A:Accession: E72124
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-158 <ARN>
A:Cross-references: GB:AE001591; GB:AE001363; NID:94376311; PIDN:AAD18213.1; PID:94377
A:Experimental source: strain CWL029
C:Genetics:
A:Gene: ptsN1
C:Superfamily: phosphotransferase system enzyme II; phosphotransferase system mannito
F:28-156/Domain: phosphotransferase system mannitol-specific enzyme II factor III hom

Query Match 1.6%; Score 7; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 LSTLTL 11
|||||
DB 126 LSTLTL 132

RESULT 11
A54506
tubulin alpha-1 chain - Plasmodium yoelii (fragment)
C:Species: Plasmodium yoelii
C:Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 13-Aug-1999
C:Accession: A54506
R:Akella, R.; Arasu, P.; Vaidya, A.B.
Mol. Biochem. Parasitol. 30, 165-174, 1988
A:Title: Molecular clones of alpha-tubulin genes of Plasmodium yoelii reveal an unusu
A:Reference number: A54506; MUID:89014607
A:Accession: A54506
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-197 <AK>
A:Cross-references: GB:M29816; NID:9160729; PIDN:AAA29779.1; PID:9160730
C:Superfamily: tubulin

Query Match 1.6%; Score 7; DB 2; Length 197;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120

Db 82 IKTKRST 88

RESULT 12

TI9438
hypothetical protein C25A1.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C/Accession: T19438

R:Morimore, B.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19124

A:Accession: T19438

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-219 <NTL>

A:Cross-references: EMBL:281038; PIDN:CAB02762.1; GSPDB:GN00019; CESP:C25A1.1

A:Experimental source: clone C25A1

C:Genetics:

A:Gene: CESP:C25A1.1

A:Map position: 1

A:Introns: 33/1; 66/1; 151/2

Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 219;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 158 KDAKOLE 164

Db 52 KDAKOLE 58

RESULT 13

C84976

tRNA (guanine-N1)-methyltransferase (EC 2.1.1.31) [imported] - Buchnera sp. (strain APS

C/Species: Buchnera sp.

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001

C/Accession: C84976

R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.

Nature 407, 81-86, 2000

A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A

A:Reference number: A84930; MUID:20445173

A:Accession: C84976

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-237 <STO>

A:Cross-references: GB:AP000398; GSPDB:GN00144

A:Experimental source: strain APS

C:Genetics:

A:Gene: trmD; BU396

C:Superfamily: tRNA (guanine-N1) methyltransferase

C:Keywords: methyltransferase; S-adenosylmethionine

Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 237;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LAIQOAK 25

Db 78 LAIQOAK 84

RESULT 14

A61544

tubulin alpha chain - common octopus (fragment)

C/Species: Octopus vulgaris (common octopus)

C/Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 10-Jul-1998

C/Accession: A61544

R:Zinov'eva, R.D.; Aleinikova, K.S.; Tomarev, S.I.

Dokl. Akad. Nauk SSSR 302, 462-467, 1988

A:Title: Isolation and structural characterization of cDNAs coding for alpha-tubulin

A:Reference number: A61544

A:Accession: A61544

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-240 <ZIN>

C:Superfamily: tubulin

F:239-240/Cleavage site: Glu-Tyr (tubulin-specific carboxypeptidase) #status predicte

F:239-240/Cross-link: peptide (Glu-Tyr) (by tubulin-tyrosine ligase) #status predicte

Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 240;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120

Db 124 IKTKRSI 130

RESULT 15

A44959

coat protein - potato virus Y (strain N) (fragment)

C/Species: potato virus Y, PVY

C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Nov-2000

C/Accession: A44959

R:Hay, J.M.; Fellowes, A.P.; Timmerman, G.M.

Arch. Virol. 107, 111-122, 1989

A:Title: Nucleotide sequence of the coat protein gene of a necrotic strain of potato

A:Reference number: A44959; MUID:90025790

A:Accession: A44959

A>Status: preliminary

A:Molecule type: genomic RNA

A:Residues: 1-264 <HAY>

A:Cross-references: GB:M22470; NID:9333508; PIDN:AAA47182.1; PID:9333509

C:Superfamily: tobacco etch virus genome polypeptide

Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 264;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 TKKDAQ 162

Db 8 TKKDAQ 14

Search completed: May 9, 2002, 12:56:34
Job time: 604 sec

This Page Blank (uspto)

QY 53 VQSATQSASTDTANPLDEHEPELYTTALENKMLINCASALNODIMRLACYDTLVHGEP 112
|||||
Db 53 vgsatqsastdtanpldehepeylttalenkmlincsalngdlmracydtlvhgetpa 112
|||||
QY 113 VIKTKRSIRLDETITWQTIKGPQVVOETDPFELMGNEKGM.LTKKDAKOLEYAAKOFTP 172
|||||
Db 113 viktksirldetitwqtikgpqvvygetcdpflimgnekgm.ltkkdaqlayaakqftp 172
|||||
QY 173 LSLSPDLDRNRP.LWSSRP.HNPMYVLP.IFMHGKPNRS.PNTPSHEAKOFTPEFRAPELKP 232
|||||
Db 173 lsisfldrnrnp.lwsrpnmpyvlplfmhggknrsptnpshearqfipnefrapelkf 232
|||||
QY 233 QVSVKYKAAEDLMGTPSDLMFGITQOSHWOIFNGKNSRPRRVHDYQPEIFLTQPVYSDLP 292
|||||
Db 233 qvsvkkaeedlwgtdsdlwfgylqgshwqlfngknsrplrvhdyqpeflfqpvysdlp 292
|||||
QY 293 WDGKVRMIGMGAVHHSNGESAKLSRSMNRRAYLMAGMEWKMLTVMPRIMGRIFKEGSGSOP 352
|||||
Db 293 wdgkvrmlgmgaavhshngesaklsrsmnraylmagmewkmltvmprlmgrl fkegs9sqp 352
|||||
QY 353 DDNPDLIDYVYGDVRFELYOLENKSNI.SGTVRYNPRSGKALQUDYVYPLGKGISGYFOI 412
|||||
Db 353 ddnpdlldyygygdvrflyglenknsisgtvrynprsgkalqldyyvplgkglsgyfoi 412
|||||
QY 413 FOGYGOSLIDYNNHEATSPGVGLMLNDMGL 442
|||||
Db 413 fgygysldyenneatsfgvglmndmgml 442
|||||

Search completed: May 3, 2002, 10:23:32
Job time: 253 sec

RESULT	1
AAAY85268	

[illegible]

Qy	121	RLBETWOTIGKROVYOIYOETDRIPLFMANGNKGLTKKDKAKOLEYAKOSTPRLSPDL	180
Db	121	RLDEIWLVLGKRPVLYGETCLPILFMNGEGLTKKdKdGLEyAKqfLpLSPISdId	180
Qy	181	RNNPRLWSSRPHNPVYLPIEFMHGKPNRSPNTPSHEAKOETPNBFRAPELKFOVSVKVA	240
Db	181	rnnPrlwSrphNpmyvLpIefmbgkpnrsPntpsheakqfLpnefrapELkfvsyvkva	240
Qy	241	AEDLMGTSDDLMFGTQOOSHQIENGNSRPFRRVDYOPETPLPOVYSQLPMDGKXRM	300
Db	241	aedlmgtsddlmfgtqooshqIengnsrpfrrvdyOpEtPlPoVySqlPmdgKxrmI	300
Qy	301	GMGAVHSHNGISAKLSRSMNFRATLMAGMEWKNLTVMFRIGRITFEBSGSQPDNDPILD	360
Db	301	gmGavhshngesaklsrsmnFrAtlmAgmEwknLtvMfrIgrITfeBsgsqpdndpIlD	360
Qy	361	YVYGCDVFEVLQLENNKSNISGTVEYNRPSRGALQDLVYVPLKGISIGYQIOFGVQOSL	420
Db	361	yYyGcdvFevLqLeNNkSnISgtVeYnrPsRGALqdlVYvPlKgISiGYqIOFGvQosL	420
Qy	421	IDYNHAEATSEGVGLMLNDPMGL	442
Db	421	IdYnhEaTsIgvGlmlNdpmGl	442

RESULT 2
 ID AAY85271 standard; Protein; 442 AA.
 AAY85271;
 29-JUN-2000 (first entry)
 BASB034 amino acid sequence #4.
 Moraxella catarrhalis infection; BASB034; diagnosis; staging;
 vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 hearing loss; antibacterial drug.
 Moraxella catarrhalis.
 WO200015802-A1.
 23-MAR-2000.
 14-SEP-1999; 99WO-EP06781.
 14-SEP-1998; 98GB-0020002.
 (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 Ruelle J;
 WPI: 2000-271440/23.
 N-PSDB: AAA10703.
 Novel BASB034 polynucleotides and polypeptides from Moraxella
 catarrhalis used to prepare vaccines against bacterial infections -
 Claim 3; Page 69; 106pp; English.
 This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 strain Mc2969. The invention relates to BASB034 polypeptides from
 M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
 polynucleotides and polypeptides may be employed as research reagents and
 material for the discovery of treatments and diagnostics for diseases,
 particularly human diseases. They are particularly used to diagnose and
 treat M. catarrhalis infections. They can be used for diagnosis of
 disease, staging of disease, or determining response of an infectious
 organism to drugs. The polynucleotides may be used as a source for
 hybridization probes, and for screening of genetic mutations, serotype,

organism or strain identification, identification of mutations in BASB034 sequences, and as components of arrays which are useful for diagnostic and prognostic purposes. The polypeptides can be used to produce antibodies. The polypeptides can also be used in vaccine formulations, and to identify agonists and antagonists. The polypeptides, antibodies, agonists and antagonists (which are bacteriostatic) are used for the treatment and prevention of diseases such as otitis media in infants and children, pneumonia in elderlies, sinusitis, nosocomial infections and invasive diseases, and chronic otitis media with hearing loss. The polypeptides, agonists and antagonists are also used for screening of antibacterial drugs. The BASB034 products of the invention can be used as screen for new antibacterial compounds that may target resistant bacteria.

Query Match	74.0%	Score 327	DB 21	Length 442
Best Local Similarity	99.8%	Pred. No. 0	Mismatches 1	Indels 0
Matches 427	Conservative 0	0	Gaps 0	
QY	15	CFALAI1QQAQAVNPVAFVDEVRESENDICGQDNELRIDVQATQSGASTDITANPLDENEP	74	
Db	15	cfal1a1qgaqavnpvafdevresnd1ygdhnelr1dvqsatgsactcnap1dehnepe	74	
QY	75	LYTALLENKTMILNCSALNODIMRLACDPLVHGENFPAVITKRSIRLDETWTQIKGR	134	
Db	75	lytalenktmlncsalngdlmr1acydd1vhngepavikctktr1rdeeltwql1ykr	134	
QY	135	QVIVQETTPDPIFLMGNEKGMILTKKDAKQDELYAAKQFTPLSLSPFDLDRNTPLMSSRPHN	194	
Db	135	qvvyqettdp1flmgnekgm1tkkdaql1eyakqftplslsfld1drntrplwsarhnp	194	
QY	195	MYVLRIRMHKGRPNRSPTPTSHNEAKQTPNDFAPRELKQVSYKVKAAADLNGTSDIMEG	254	
Db	195	myv1r1rlmhgrpnrsptptshneakqct1pnefaprelk1qvsykvkaaed1wgtddsl1vf9	254	
QY	255	YTOOSHQIQFNCKNSRPFRRVNDQREIFLTQPVYSD1PLPDGCVNRHIGAGVHNSGESAK	314	
Db	255	ytqgshwq1fngknsrpfrrvhdqypr1fltqpyad1pndgkvnm1gmav1hnsgeak	314	
QY	315	LSRSWNRNAYILMAGNEKMLTVMPRIRNGRIFKESSGQPRDNDLIDYGYGCVRR1YQLE	374	
Db	315	ls1rswnrnayilmagnekmltvmpr1wgr1fk1egsgsqpdpnd1ldyy9ygdvrr1yql1e	374	
QY	375	NKNSIGTVRYNPRSGKALQLDYVYPLGKISGVEQIQGQSGSLIDYNHEDATSFVGVL	434	
Db	375	nkn1sigtvyrynprrsgkalqldyvypl1gk1sgyfq1e1qy9gsl1dyhneatsf1ygl1	434	
QY	435	MLNDPMGL 442		
Db	435	mlndpmgl 442		
RESULT 3				
AAy85269				
ID	AAy85269 standard; Protein; 442 AA.			
XX	AAy85269;			
XX	29-JUN-2000 (first entry)			
XX	BASB034 amino acid sequence #2.			
XX	Moraxella catarrhalis infection; BASB034; diagnosis; staging;			
XX	vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;			
XX	sinusitis; nosocomial infection; invasive disease; chronic otitis media;			
XX	hearing loss; antibacterial drug.			
XX	Moraxella catarrhalis			
XX	WO200015802-A1.			

Query Match	Best Local Similarity	Score	DB	Length
Matches 224: Conservative	100.0%	50.7%	224	442
			Pred. No. 9.2e-209	
			Mismatches 0: Indels 0: Gaps 0	
219 OPTPEFRARPELKFQVSVAVKAADIMGSDIMFEGITQOSHMQITNGKSRFRVNDYQ	278			
219 qftpefrarpeLkfQvsVvkaaedtwgdsdtdwfgtqshwqfngknsrpfvrvhdyq	278			
279 PEIPLTOPRYSLLPMDGKVRMTGMAVNHNSGSAATLSWNAAYLMAAGEMKNTLTPMR	338			
279 peiLltpQrYsdlpWdgkVrMtgmaVnhnsGesaLtswnaYlmaagewMknltpmr	338			
339 IWGRIFKEGSGSQPDNDPDLIDYGYGADVRFYQLEKNKSNISGTVRYNPRSGKALOLDY	398			
339 IwgrIfkEgsgsqpDndpDlIdyGyGadvRfYqLeKnKsnISgTvRYnPrsGkaLoldY	398			
399 VYPLGKGISGYQIFOGYQSLIDYVNHKATSFEGVGLMLNDWML	442			
399 vYpLgKgisgyQifogYqSliDyVnhKaTsfegVglMlNdWmL	442			

XX	BASB034 amino acid sequence #3.
DE	
XX	Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KW	vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KM	sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KV	hearing loss; antibacterial drug.
XX	
OS	.Moraxella catarrhalis.
XX	
PN	WO200015802-A1.
XX	
PD	23-MAR-2000.
XX	
PF	14-SEP-1999; 99WO-EP06781.
XX	
PR	14-SEP-1998; 98GB-0020002.
XX	
PA	(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX	
PI	Ruelle J;
XX	
DR	WPI; 2000-271440/23.
DR	N-PSDB: AAA10702.
XX	
PT	Novel BASB034 polynucleotides and polypeptides from Moraxella
XX	catarrhalis used to prepare vaccines against bacterial infections -
PS	Claim 3; Page 68; 106pp; English.
XX	
CC	This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC	strain Mc2913. The invention relates to BASB034 polypeptides from
CC	M. catarrhalis strains Mc2911, Mc2908, Mc2913 and Mc2969. The BASB034
CC	polynucleotides and polypeptides may be employed as research reagents and
CC	material for the discovery of treatments and diagnostics for diseases,
CC	particularly human diseases. They are particularly used to diagnose and
CC	treat M. catarrhalis infections. They can be used for diagnosis of
CC	disease, staging of disease, or determining response of an infectious
CC	organism to drugs. The polynucleotides may be used as a source for
CC	hybridization probes, and for screening of genetic mutations, serotype,
CC	organism or strain identification, identification of mutations in BASB034
CC	sequences, and as components of arrays which are useful for diagnostic
CC	and prognostic purposes. The polypeptides can be used to produce
CC	antibodies. The polypeptides can also be used in vaccine formulations,
CC	and to identify agonists and antagonists. The polypeptides, antibodies,
CC	agonists and antagonists (which are bacteriostatic) are used for the
CC	treatment and prevention of diseases such as otitis media in infants and
CC	children, pneumonia in elderlies, sinusitis, nosocomial infections and
CC	invasive diseases, and chronic otitis media with hearing loss. The
CC	polypeptides, agonists and antagonists are also used for screening of
CC	antibacterial drugs. The BASB034 products of the invention can be used
CC	screen for new antibacterial compounds that may target resistant
CC	bacteria.
CC	
XX	
SQ	Sequence 442 AA:
XX	
Query Match	50.7%; Score 224; DB 21; Length 442;
Best Local Similarity	100.0%, Pred. No. 9,2e-209;
Matches 224; Conservative	0; Mismatches 0; Indels 0; Gaps 0
OY	GFTPNEFRAPELKFQVSVKVKAEDIMGTDSDLMFGYTQOOSHMQIFNGKNSRPFRYHDYQ 278
Db	
219	gftpefrapelfkfgsvvkkaedltwgsdsltlfyqtqshmqifngksrpfryhdxy 278
OY	PETFLTOPYSPDLPMDGKRMTGMGAHVHSHNSESALSTSWNRATYLMAEMENKNTLVMR 338
Db	
279	petfltpypysdlpwogkrmtgmgaavhshnsesalstswmraylmaqemwnltvmpr 338
OY	IGRGIFKEGSGSQDPDNPIILDYYGYGVDFLYOLENKSNISGTAVYNFRSGGALQLDY 398
Db	
339	ivgrifkegsqsqpdpndpiildyygydvrtfilyqlenksnisgtvyrnpsrkgalqldy 398

Thu May 9 12:27:06 2002

us-09-787-083-2.oli.rag

Page 4

Qy 399 VYDLGKISGTFQIQFGIGYQSLIDYNHEATSFVGVLMDNMKL 442
|||||
Dd 399 VYPLGKIGSYFQIFGYYQSLIDYNHEATSFYGVLMINDMGI 442

Search completed: May 3, 2002, 10:23:31
Job time: 252 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2002, 10:23:31 ; Search time 71.76 Seconds

(Without alignments)
456.249 Million cell updates/sec

Title: US-09-787-083-4

Perfect score: 442
Sequence: 1 MKVSLSTLTLSILPCEFAIIA.....YHNEATSPGVGIMLNDMGL 442Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 522463 seqs, 74073290 residues

Word size : 15

Total number of hits satisfying chosen parameters: 4

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

A_Geneseq_1101:*

1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	442	100.0	442	21	AAV85269 BASB034 amino acid
2	390	88.2	442	21	AAV85270 BASB034 amino acid
3	289	65.4	442	21	AAV85271 BASB034 amino acid
4	224	50.7	442	21	AAV85268 BASB034 amino acid

ALIGNMENTS

RESULT 1
AAV85269

ID	AAV85269 standard; Protein: 442 AA.
XX	
AC	AAV85269;
XX	
DT	29-JUN-2000 (first entry)
XX	
DE	BASB034 amino acid sequence #2.
XX	
KW	Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KW	vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW	sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW	hearing loss; antibacterial drug.
XX	
OS	Moraxella catarrhalis.
XX	
PN	MO200015802-A1.
PD	
XX	23-MAR-2000.
XX	
PF	14-SEP-1999; 99WO-EP06781.
XX	
PR	14-SEP-1998; 98GB-0020002.
XX	
PA	(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX	
PI	Ruelle J;
XX	
DR	WPI; 2000-271440/23.
DR	N-PSDB; AAA10701.
XX	
PT	Novel BASB034 polynucleotides and polypeptides from Moraxella
PT	catarrhalis used to prepare vaccines against bacterial infections
XX	
XX	Claim 3; Page 67; 106pp: English.
XX	
CC	This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC	strain Mc2908. The invention relates to BASB034 polypeptides from
CC	M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
CC	polynucleotides and polypeptides may be employed as research reagents and
CC	material for the discovery of treatments and diagnostics for diseases,
CC	particularly human diseases. They are particularly used to diagnose and
CC	treat M. catarrhalis infections. They can be used for diagnosis of
CC	disease, staging of disease, or determining response of an infectious
CC	organism to drugs. The polynucleotides may be used as a source for
CC	hybridization probes, and for screening of genetic mutations, serotype,
CC	organism or strain identification, identification of mutations in BASB034
CC	sequences, and as components of arrays which are useful for diagnostic
CC	and prognostic purposes. The polypeptides can be used to produce
CC	antibodies. The polypeptides can also be used in vaccine formulations,
CC	and to identify agonists and antagonists. The polypeptides, antibodies,
CC	agonists and antagonists (which are bacteriostatic) are used for the
CC	treatment and prevention of diseases such as otitis media in infants and
CC	children, pneumonia in elderlies, sinusitis, nosocomial infections and
CC	invasive diseases, and chronic otitis media with hearing loss. The
CC	polypeptides, agonists and antagonists are also used for screening of
CC	antibacterial drugs. The BASB034 products of the invention can be used
CC	screen for new antibacterial compounds that may target resistant
CC	bacteria.
XX	
XX	
XX	
SO	Sequence 442 AA;
XX	
Query Match	100.0%; Score 442; DB 21; Length 442;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 MKVSLSTLTLSILPCEFAIIAQAAPVAFVDEVRKNDLGDNDLIGVOSATQSA 60
DB	1 mkvslstltlslpcefallaiaqgavpafvdevrskndlgdneillgvsaqsa 60
OY	61 STDTPANPLDEHEPELYTALNKTMLINCSALNDIMRLACDYLHGEFPAVVKTKRSI 120
DB	61 stdtcanpldehepeelytalnktmlincsalndimrlacdydlvhgetpavvktkrsi 120

PR 14-SEP-1998; 98GB-0020002.
 XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 PA Ruelle J;
 XX WPI: 2000-2714A0/23.
 DR N-PSDB; AAA10703.
 XX Novel BASB034 polynucleotides and polypeptides from Moraxella
 PT catarrhalis used to prepare vaccines against bacterial infections
 PS Claim 3; Page 69; 106pp; English.
 XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 CC strain Mc2969. The invention relates to BASB034 polypeptides from
 CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
 CC polynucleotides and polypeptides may be employed as research reagents and
 CC material for the discovery of treatments and diagnostics for diseases,
 CC particularly human diseases. They are particularly used to diagnose and
 CC treat M. catarrhalis infections. They can be used for diagnosis of
 CC disease, staging of disease, or determining response of an infectious
 CC organism to drugs. The polynucleotides may be used as a source for
 CC hybridization probes, and for screening of genetic mutations, serotype,
 CC organism or strain identification, identification of mutations in BASB034
 CC sequences, and as components of arrays which are useful for diagnostic
 CC and prognostic purposes. The polypeptides can be used to produce
 CC antibodies. The polypeptides can also be used in vaccine formulations,
 CC and to identify agonists and antagonists. The polypeptides, antibodies,
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.
 XX Sequence 442 AA;
 SQ

Query Match 65.4%; Score 289; DB 21; Length 442;
 Best Local Similarity 99.7%; Pred. No. 7.6e-273;
 Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 53 VQASQTASDTNPANPLDNEHEPELYTALENKTMILNCSALNODIMRLACYDTLVHGETPA 112
 DB 53 vgsatgsastltanpldehepeyltalenktmlncsalngdmlrlacydtlvhgetpa 112
 QY 113 VIKRKRSIRLDETWOTIKGKPOVYVOETDPIFLMGNEKGMLEKKAQOLEVAAKQFTP 172
 DB 113 vikrkrsirldetwotikgkpoovyvoetdpriflmgnekgmlekkakqoleyaakqftt 172
 QY 173 LLSFEDDRNTPLMSSRNPNRYVLPPIFMHGKPNRSPNPSHAROPTPEFAPELKFE 232
 DB 173 llsfddidmrtplwsrtnpnmryvlpifmhgkpnrsprnpsneakqftrpefapekfkf 232
 QY 233 QVSVKVAKAEDLMGTSDDLWFGYTQOSHWOIFNGKNSRPFVNDYOEPIFLTOPVSDLP 292
 DB 233 qvsavkvaeeadlwgtddslwfgytcqgshwqifngknsrpfvndhyoeifltqpvssdlp 292
 QY 293 WDGVKVRMIGMKAHVHNSGESAKLSRSWNKRAYLMAGMEKMLTVNPRIMGRIFKSGSGSQP 352
 DB 293 wdgvkvrmlgmgaavhnsgeaklsrswntaylmagmekmltvmprilwgrlffksgsgsqp 352
 QY 353 DDNDPDLIDYGYGVDFEYOLENKSNSGTVRPNRSGKALODYVYPLGKGSIGYFOI 412
 DB 353 ddndpdlidygygvdfeyoleknsnsgtvrpnrsqkalgldyvyplgkgsigyfoi 412
 QY 413 FQGYGQSLIDYNNHEATSFVGVLMDMMGL 442
 DB 413 fggygqslidyynheatsfvgvglmmdmmgl 442

RESULT 4
 ID AAY85268 standard; Protein; 442 AA.
 XX
 AC AAY85268;
 XX
 DT 29-JUN-2000 (first entry)
 XX
 DE BASB034 amino acid sequence #1.
 XX
 KW Moraxella catarrhalis infection; BASB034; diagnosis; staging;
 KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KW hearing loss; antibacterial drug.
 XX
 OS Moraxella catarrhalis.
 XX
 PN M0200015802-A1.
 XX
 PD 23-MAR-2000.
 XX
 PF 14-SEP-1999; 99WO-EP06781.
 XX
 PR 14-SEP-1998; 98GB-0020002.
 XX
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 PI Ruelle J;
 XX WPI: 2000-2714A0/23.
 DR N-PSDB; AAA10700.
 XX
 PT Novel BASB034 polynucleotides and polypeptides from Moraxella
 PT catarrhalis used to prepare vaccines against bacterial infections
 PS Claim 3; Fig 2; 106pp; English.
 XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 CC strain Mc2931 (ATCC 43617). The invention relates to BASB034 polypeptides
 CC from M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The
 CC BASB034 polynucleotides and polypeptides may be employed as research
 CC reagents and material for the discovery of treatments and diagnostics for
 CC diseases, particularly human diseases. They are particularly used to
 CC diagnose and treat M. catarrhalis infections. They can be used for
 CC diagnosis of disease, staging of disease, or determining response of an
 CC infectious organism to drugs. The polynucleotides may be used as a source
 CC for hybridization probes, and for screening of genetic mutations,
 CC serotype, organism or strain identification, identification of mutations
 CC in BASB034 sequences, and as components of arrays which are useful for
 CC diagnostic and prognostic purposes. The polypeptides can be used to
 CC produce antibodies. The polypeptides can also be used in vaccine
 CC formulations, and to identify agonists and antagonists. The polypeptides,
 CC antibodies, agonists and antagonists (which are bacteriostatic) are used
 CC for the treatment and prevention of diseases such as otitis media in
 CC infants and children, pneumonia in elderlies, sinusitis, nosocomial
 CC infections and invasive diseases, and chronic otitis media with hearing
 CC loss. The polypeptides, agonists and antagonists are also used for
 CC screening of antibacterial drugs. The BASB034 products of the invention
 CC can be used screen for new antibacterial compounds that may target
 CC resistant bacteria.
 XX Sequence 442 AA;
 SQ

Query Match 50.7%; Score 224; DB 21; Length 442;
 Best Local Similarity 100.0%; Pred. No. 1.4e-209;
 Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 219 QFTPNFRAPELKFOVSVKVAKAEDLMGTSDDLWFGYTQOSHWOIFNGKNSRPFVNDYQ 278
 DB 219 qftpnfrapelkfovsvkvaeeadlwgtddslwfgytcqgshwqifngknsrpfvndyq 278

```

QY 279 PEIFLQPVYSDLPMQCKVBMIGMAVHSHNGESAKLSRSWNRAIYMAGHEWKLTVMPR 338
    |||
Db 279 pelfltpvysdipwdgkxvmlgmaavhshngesaklsrswrnylmaamewknltvmpr 338
    |||
QY 339 IMGRIFKEGSGSQPDDNPDILDYGYGQDVRFLYOLENKSNSIGTVRYRNPNSGKGALQLDY 398
    |||
Db 339 lwgrlftkgsqsgpddnplldyygydvrflyqlenksnlsqclvrynpnsqkgaqlldy 398
    |||
QY 399 VYPLGKGISGYFOIFQGYGQSLIDYNHEATSFGVGLMLNDMGL 442
    |||
Db 399 vylpgkgsygyfqlfgygqslldynheatsfgvglmndmgl 442
    |||

```

Search completed: May 3, 2002, 10:23:31
 Job time: 252 sec

PR 14-SEP-1998: 98GB-0020002.
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX Ruelle J;
PI WPI: 2000-271440/23.
DR N-PSDB; AAA10703.
XX
PT Novel BASB034 polynucleotides and polypeptides from Moraxella
PT catarrhalis used to prepare vaccines against bacterial infections
XX
PS Claim 3; Page 69; 106pp: English.
XX
CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2969. The invention relates to BASB034 polypeptides from
CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
CC polynucleotides and polypeptides may be employed as research reagents and
CC material for the discovery of treatments and diagnostics for diseases,
CC particularly human diseases. They are particularly used to diagnose and
CC treat M. catarrhalis infections. They can be used for diagnosis of
CC disease, staging of disease, or determining response of an infectious
CC organism to drugs. The polynucleotides may be used as a source for
CC hybridization probes, and for screening of genetic mutations, serotype,
CC organism or strain identification, identification of mutations in BASB034
CC sequences, and as components of arrays which are useful for diagnostic
CC and prognostic purposes. The polypeptides can be used to produce
CC antibodies. The polypeptides can also be used in vaccine formulations,
CC and to identify agonists and antagonists. The polypeptides, antibodies,
CC agonists and antagonists (which are bacteriostatic) are used for the
CC treatment and prevention of diseases such as otitis media in infants and
CC children, pneumonia in elderly, sinusitis, nosocomial infections and
CC invasive diseases, and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.
XX
XX Sequence 442 AA:
SO
Query Match 71.5%; Score 316; DB 21; Length 442;
Best Local Similarity 99.8%; Pred. No. 3.6e-299;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 26 AVPNPVADEVYRSENDLGQNEPLIDVQASATQASTDPANFLDEHEPELYTTALENKTW 85
DB 26 avnpvatdevyrsendlgqndelpidvqasatqastdpanldehepeyltalentkm 85
QY 86 LINCASALNODIMRLACQYTLVHGETPANYIKTRSRILDETTMOTIKGAPVYVQETTPDI 145
DB 86 lincasalnqdimrlacqytlvhgetpanyiktrsrildetltmqtkgprvyvqettpdi 145
QY 146 FLMGSEKGLTFFKKDAKOLEYAAKQFTPLSLFPLDRNNTPLMSSRPHPMVYLPIFFMKGK 205
DB 146 flmgsekmltkkdaakqleyaakqftplslsfldlrmncplwssrphpmvylpifmngk 205
QY 206 PNRSPNTPSHEARQFTPNFEFAPELKFOVSVKVAEEDLWGTDSULMGYTOQSHMOJFN 265
DB 206 pnrspntrpshearkftpnfefrapelkfgvsvkvaeadlwtgdsdlwtgysqshwqjfn 265
QY 266 GKNSRPFVHYOYPERIFLTQPYSDLPMDGKYRMTGKAGVAHNSGSAKLSMNRATYIM 325
DB 266 gknsrpfvhydypei fltqpyssdlpmdgkvrmtgkagvahnsgesaklsrswratyim 325
QY 326 AGMEKNLTVMRIRMGRIFFKEGSGSQPDNDPILIDYGYGVDRFLYQLEKNKSNIGTVRY 385
DB 326 agmeknltvmprlmgri fffkegsgsqpddndpildygygvdrflfyqlenksnigtvry 385
QY 386 NPSGKALQLDYVYPLCKGISGYFOIFQGYGQSLIDYNHEATSFVGLMLNDMWGL 442
DB 386 npsgkalgldyvyplckgisgyfoifqgygqslidyneatsfvglmlndmwgl 442

RESULT 4
AAV85268
ID AAV85268 standard; Protein: 442 AA.
XX
XX AAV85268;
AC
XX 29-JUN-2000 (first entry)
DT
XX
XX BASB034 amino acid sequence #1.
DE
XX
KW Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW hearing loss; antibacterial drug.
XX
OS Moraxella catarrhalis.
XX
XX WO200015802-A1.
XX
XX 23-MAR-2000.
XX
XX 14-SEP-1999; 99WO-EP06781.
XX
XX 14-SEP-1998; 98GB-0020002.
XX
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Ruelle J;
XX
XX WPI: 2000-271440/23.
XX
XX N-PSDB; AAA10700.
XX
PT Novel BASB034 polynucleotides and polypeptides from Moraxella
PT catarrhalis used to prepare vaccines against bacterial infections
XX
PS Claim 3; Fig 2; 106pp; English.
XX
CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2931 (ATCC 43617). The invention relates to BASB034 polypeptides
CC from M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The
CC BASB034 polynucleotides and polypeptides may be employed as research
CC reagents and material for the discovery of treatments and diagnostics for
CC diseases, particularly human diseases. They are particularly used to
CC diagnose and treat M. catarrhalis infections. They can be used for
CC diagnosis of disease, staging of disease, or determining response of an
CC infectious organism to drugs. The polynucleotides may be used as a source
CC for hybridization probes, and for screening of genetic mutations,
CC serotype, organism or strain identification, identification of mutations
CC in BASB034 sequences, and as components of arrays which are useful for
CC diagnostic and prognostic purposes. The polypeptides can be used to
CC produce antibodies. The polypeptides can also be used in vaccine
CC formulations, and to identify agonists and antagonists. The polypeptides,
CC antibodies, agonists and antagonists (which are bacteriostatic) are used
CC for the treatment and prevention of diseases such as otitis media in
CC infants and children, pneumonia in elderly, sinusitis, nosocomial
CC infections and invasive diseases, and chronic otitis media with hearing
CC loss. The polypeptides, agonists and antagonists are also used for
CC screening of antibacterial drugs. The BASB034 products of the invention
CC can be used screen for new antibacterial compounds that may target
CC resistant bacteria.
XX
XX Sequence 442 AA:
SO
Query Match 50.7%; Score 224; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 6.6e-209;
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 219 QFTPNFEFAPELKFOVSVKVAEEDLWGTDSULMGYTOQSHMOJFNKNSRPFVHYQ 278
DB 219 qftpnfefrapelkfgvsvkvaeadlwtgdsdlwtgysqshwqjfnknsrpfvhyq 278

OY 279 PEIPLTQPVYSDLPWDGKVRNIGMGAVHHSNGESAKLSRSNRAVYLMAGMEKNLTVMPR 338
Db 279 PEIPLTQPVYSDLPWDGKVRNIGMGAVHHSNGESAKLSRSNRAVYLMAGMEKNLTVMPR 338
OY 339 IWGRIFKEGSGSQPPDDNPDLIDYXGYGDVRFLYOLENKSNI SGTVRYNPRSGKALOLDY 398
Db 339 Iwgrlfkegsgsqppddnpdlldyygydvrflyqlenksnls gtvrynprrsgkaloldy 398
OY 399 VYPLGKGISGYFQJFQGYGQSLIDYNHEATSEGVGLMLNDMMGL 442
Db 399 vyplgkgisgyfqjfggygqslldynheatsfyvglmndmmgl 442

Search completed: May 3, 2002, 10:23:31
Job time: 252 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2002, 10:24:14 ; Search time 36.69 Seconds
(without alignments)
271.094 Million cell updates/sec

Title: US-09-787-083-6

Perfect score: 442
Sequence: 1 MKVSLSTLTLSLSCPAIIA.....YNHEATSFQVGLMDWGL 442

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 212252 seqs, 22503292 residues

Word size : 15

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents.AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/1aa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description

No matches found

Search completed: May 3, 2002, 10:24:14
Job time: 185 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:26:28 ; Search time 77.9 Seconds
(without alignments)
829.941 Million cell updates/sec

Title: US-09-787-083-8
Perfect score: 2363
Sequence: 1 MKVSLSTLTSLPCFALIA.....YNHETSRGVLMLNDWGL 442

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 segs, 146272329 residues
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP:REMBL_17:*
2: SP:archaea:*
3: SP:bacteria:*
4: SP:fungi:*
5: SP:human:*
6: SP:invertebrate:*
7: SP:mammal:*
8: SP:organelle:*
9: SP:phage:*
10: SP:plant:*
11: SP:rodent:*
12: SP:virus:*
13: SP:vertebrate:*
14: SP:unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	815	34.5	382	2	09K0U7
2	815	34.5	409	2	09J721
3	372	15.7	306	2	09CL22
4	366	15.5	329	2	09PMU8
5	360.5	15.3	292	2	09SID7
6	360	15.2	289	2	09LE6N
7	356.5	15.1	292	2	09Z4N8
8	342.5	14.5	297	2	032349
9	246.5	10.4	355	2	025241
10	240.5	10.2	355	2	09ZLX5
11	123	5.2	278	2	09XB53
12	117	5.0	1686	4	000443
13	111.5	4.7	602	11	063485
14	108.5	4.6	821	2	059241
15	107.5	4.5	824	2	09F216
16	106.5	4.5	901	5	018749
17	106	4.5	798	4	09UN32
18	106	4.5	798	4	09UBK2
19	103	4.4	435	2	09X9C0

20	102.5	4.3	739	5	09V9E6	09V9E6 drosophila
21	102.5	4.3	1046	2	084941	084941 streptococc
22	101.5	4.3	403	5	P91736	P91736 hydra magni
23	101.5	4.3	783	2	043554	043554 bacillus sp
24	100.5	4.3	660	2	09K6K8	09K6K8 bacillus ha
25	100.5	4.3	1509	11	061194	061194 mus musculu
26	100.5	4.3	1658	11	061182	061182 mus musculu
27	100.5	4.3	3247	12	065553	065553 bovine herp
28	99.5	4.2	797	2	09RDM6	09RDM6 lactobacill
29	99	4.2	422	2	09RCZ6	09RCZ6 streptomyc
30	99	4.2	5005	2	09P2P5	09P2P5 ureaplasma
31	98.5	4.2	467	4	09UJY5	09UJY5 homo sapien
32	98.5	4.2	467	4	09URY2	09URY2 homo sapien
33	98.5	4.2	788	10	09SCV5	09SCV5 arabidopsis
34	98.5	4.2	838	10	09C8I5	09C8I5 arabidopsis
35	97.5	4.1	683	2	09A6U7	09A6U7 caulobacter
36	96.5	4.1	749	2	059154	059154 anaerocellu
37	96	4.1	791	10	049137	049137 arabidopsis
38	96	4.1	4307	5	019319	019319 caenorhabdi
39	95.5	4.0	576	2	09L1I5	09L1I5 streptomyc
40	95.5	4.0	772	2	09Z3V2	09Z3V2 pseudomonas
41	95.5	4.0	1372	10	09FL92	09FL92 arabidopsis
42	95	4.0	523	4	09H856	09H856 homo sapien
43	95	4.0	765	2	054183	054183 streptomyc
44	95	4.0	871	4	09Y5C3	09Y5C3 homo sapien
45	95	4.0	938	4	09Y5F7	09Y5F7 homo sapien

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	382 AA.
09K0U7	09K0U7			
AC	09K0U7			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	PHOSPHOLIPASE A1, PUTATIVE.			
GN	NMB0464.			
OS	Neisseria meningitidis (serogroup B).			
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.			
OX	NCBI_TaxID=491;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MC58 / SEROGROUP B;			
RX	MEDLINE=20175755; PubMed=10710307;			
RA	Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,			
RA	Eisen J.A., Ketchum K.A., Hood D.W., Feden J.F., Dodson R.J.,			
RA	Nelson W.C., Gwinn M.W., Deboy R., Peterson J.D., Hickey E.K.,			
RA	Hait D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,			
RA	Mason T., Ciecko A., Parksey D.S., Blair E., Cifton H., Clark E.B.,			
RA	Collon M.D., Ullrichback T.R., Khouri H., Qin H., Yamathavan J.,			
RA	Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,			
RA	Smith H.O., Fraser C.M., Moxon E.R., Rappelli R., Venter J.C.,			
RT	"Complete genome sequence of Neisseria meningitidis serogroup B strain			
RT	MC58."			
RL	Science 287:1809-1815(2000).			
DR	EMBL: AE002403; AAF40901.1; -			
DR	TIQR: NMB0464; -			
DR	InterPro: IPR003197; PLAI.			
DR	Pfam: PF02253; PLAI: 1.			
KW	Complete proteome.			
SO	SEQUENCE 382 AA; 42714 MW; B468A802F062E836 CRC64;			

Query Match 34.5%; Score 815; DB 2; Length 382;
Best Local Similarity 44.7%; Pred. No. 8.7e-63;
Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;
QY 87 INCSALNODMIRLACYDTLVHGETPAVI-----KTKSRIRLDETIVQFI-KGRPVVYQE 140
DB 33 LQCAALIDNVTIRLACYRIRFAOQLPSSAGOGESKAVLNLTEIVRSRLDKGEAVIYVEK 92

Qy	141	TTDPIFLMGNEKGLTKKDKKQLEAYAKOCTPILSDLDLRN-TPMSSRPNNPVLP	199
Db	93	GGDL-----PADSAGETADITYPLISMATDLDKNDRLGLGVRHNPTLMP	139
Qy	200	IFMHGKPNRSNPESH-EAKOFTPNERABELKFOYSVKVAAEDLMGTDSDLMFGTQO	258
Db	140	LMYNNSPVYAGSPTRGTQVQEKFGQCKRAETKLQVSEFKSIADLEFKTRADLMFGTQR	199
Qy	259	SHMOIFN-GKRSRPFRRVNDQPEFLTPQPVYSDLPMDGKVRMTGMCVAHNSNESAKLSR	317
Db	200	SDMOIYNGRRSAFPRFTDYKPELFTLPQPKADLPFGGRIRMAGAGVHOSNQSPRESR	259
Qy	318	SMNRAYLMAGEMENLTMVPRIMGRIFKEGSSQSDPNPILDYGYGDVRFPLEJENKS	377
Db	260	SMNRITMAAGHEMKGLVYIPRVYWRADQ-SCDK-NQNPILADYMGGDYKLOLYRLNDQ	317
Qy	378	NISGTVRNPRRSGKALQLDVYPLGKIGISGYRQIFOGYQSLDYNHEATSEFGVGLMN	437
Db	318	MYSLRYLRNPRTGCAIEAATFPIKGLKGVVGFHGYGESLIDYNNHKGNGIGIGIMFN	377
Qy	438	DWKG 442	
Db	378	DLDGI 382	
RESULT	2		
Q9J721			
AC	Q9J721	PRELIMINARY:	PRT: 409 AA.
DT	01-OCT-2000 (TREMBLrel, 15, Created)		
DT	01-OCT-2000 (TREMBLrel, 15, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel, 17, Last annotation update)		
DE	PUTATIVE PHOPHOLIPASE.		
GN	NMA2021.		
OS	Neisseria meningitidis (serogroup A).		
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.		
OX	NCBI_TaxID=65699;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-22491 / SEROGROUP A / SEROTYPE 4A;		
RX	MEDLINE-20222556; PubMed-10761919;		
RA	Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,		
RA	Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,		
RA	Davies R.M., Davis P., Devlin K., Feltham T., Hamlin N., Holtroyd S.,		
RA	Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,		
RA	Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,		
RA	Whitehead S., Spratt B.G., Barrrell B.G.;		
RT	"Complete DNA sequence of a serogroup A strain of Neisseria		
RT	meningitidis 22491."		
RL	Nature 404:502-506(2000).		
DR	EMBL: AL162757; CAB85240.1; -		
DR	InterPro: IPR003187; Pfam1.		
DR	Pfam: PF02253; Pfam1.		
KW	Complete proteome.		
SO	SEQUENCE 409 AA; 45862 MW; CD6585B064D01A41 CRC64;		

[illegible]

OY	259	SHMOEN-GKSRPFRHYDQPELEFLQRPVSDPMPGKVMIGMGAVHNSNGESAKLR	3147
		: : : : : : : : : : : : : : : : : : : : : : : : :	
Db	227	SDMQITNQGSRSAFPRNTDYKPELFLQPVAKADLPFGGRLLMGLAGVHNSNGSREPESR	286
OY	318	SMNRATYIAGMEKMLNLTVMPIRMGRIFKEGSGSQPDNDPILDYGYGDVRFYQLENKS	3777
		: : : : : : : : : : : : : : : : : : : : : : : : :	
Db	287	SMNRATYIAGMEKMLNLTVMPIRMGRIFKEGSGSQPDNDPILDYGYGDVRFYQLENKS	3444
OY	378	NISGIVRARNPSSGKALQDLYVYPLKKGICSGYFQIPGCGOGLIDYVHHEATSPGVMNLN	437
		: : : : : : : : : : : : : : : : : : : : : : : : :	
Db	345	NVSYLVRNPKXTGIGALTAATFTPIKGLKGVKRGHFGESLIDYVHNGKNGITGILMFN	404
OY	438	DMMGL 442	
		: : : : : : : : : : : : : : : : : : : : : : : : :	
Db	405	DLDGI 409	
RESULT	3		
O9CL22			
ID	O9CL22	PRELIMINARY;	PRT; 306 AA.
AC	O9CL22		
DT	01-JUN-2001 (Tremblrel. 17, Created)		
DT	01-JUN-2001 (Tremblrel. 17, Last sequence update)		
DT	01-JUN-2001 (Tremblrel. 17, Last annotation update)		
DE	HYPOTHEICAL PROTEIN PM1426.		
GN	PM1426.		
OS	Pasteurella multocida.		
OC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;		
OC	Pasteurella.		
OX	NCBI_Taxid=747;		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RC	STRAIN=PM70;		
RX	MEBLINE=21145866; PubMed=11248100;		
RA	May B.J., Zhang Q., Li L.L., Pauslian M.L., Whittam T.S., Kapur V.;		
RT	"Complete genomic sequence of Pasteurella multocida pm70.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).		
DR	EMBL: AE006179; AAK03510.1; -		
DR	InterPro: IPR003187; Pfam: 1.		
DR	Pfam: PF02253; Pfam: 1.		
SW	Hypothetical protein; Complete proteome.		
SW	SEQUENCE 306 AA; 35580 MW; EAF3DE8C1C22B26E CRC64;		

Query Match	15.78;	Score 372;	DB 2;	Length 306;
Best Local Similarity	39.06;	Pred. No. 2.5e-24;		
Matches 83;	Conservative 40;	Mismatches 82;	Indels 8;	Gaps
Oy	229	ELKEFQSVKVAAEEDLMDGTSDLMFQYQOOSHMOIFGNKNSRPERVHDYQEIPLTPQVY	288	
		::: : : : : :		
Db	97	EKEFQISALPLMKRCITLNNNSVLAASYQKSWPQSLNVDDSPRETNIEFQFLANKTQ	156	
Oy	289	SDLPMDDGKVRMIGMGAVHSNG--ESAKLSRSNNRAYLMAEMENLTVMPRIWRIFFE	346	
		::: : : : : : : : :		
Db	157	YSLPFGMTLDQVETCININQSGNRDAAEKLSSNNRFLYRAAIIQNTVETKPPMRIRPEK	216	
Oy	347	GSGSGPDDNPILDYDYGDVRF-LYOLEKSNISGTVRYNPRSGKALQLDLVYVPLGKG	405	
		::: : : : : : : : :		
Db	217	--AKNDNNPDITKRGHFDVALGGYYVHDHQFKLSG--HYNPINSGGLEASYSYPTKN	271	
Oy	406	ISGTFQIFQIGQSLIDYNHEATSEFGVGLMND	438	
		: : : : : :		
Db	272	IRFYQTYNGGESPILDYQQRIRQIRIGIGISLNN	304	
RESULT	4			
O9PM08				
O9PM08		PRELIMINARY;	PRT;	329 AA.
AC	O9PM08:			
DT	01-OCT-2000	(TREMBLeref. 15, Created)		
DT	01-OCT-2000	(TREMBLeref. 15, Last sequence update)		
DT	01-JUN-2001	(TREMBLeref. 17, Last annotation update)		

DE PHOSPHOLIPASE A (EC 3.1.1.32).
 GN PLD4 OR C13351.
 OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
 OC Campylobacter.
 OX NCBI_TaxID=197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCCTC 11168;
 RX MEDLINE=20150912; PubMed=10688204;
 RA Parkhill J., Wren B.W., Mungall K., Kelsey J.M., Churcher C.,
 RA Besham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
 RA Jagers K., Kariyasev A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Rajandream M.A., Rutherford K.M., Van Vleet A.H.M.,
 RA Whitehead S., Barrett B.G.;
 RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
 RT reveals hypervariable sequences.";
 RL Nature 403:665-668(2000).
 DR EMBL: AL139078; CAB73778.1;
 DR InterPro: IPR003187; Pfam: 1.
 DR Pfam: PF02253; PLAI: 1.
 DR PRINTS: PR01486; PHPLIPASEA1.
 KW Complete proteome.
 SQ SEQUENCE 329 AA; 38880 MW; A32731F2B751AC44 CRC64;

Query Match 15.5%; Score 366; DB 2; Length 329;
 Best Local Similarity 33.2%; Pred. No. 9.2e-24;
 Matches 105; Conservative 48; Mismatches 125; Indels 38; Gaps 13;

QY 130 IKGKPYVYQETDPIFLMKNKMLTKDAKOLEYAAKQETPISLFDLDRNT--PLW 187
 Db 43 LKSSVLLISOBONSSQATOTONSTITKEKO-----DSRLALNLYGENSEPNLG 96
 QY 188 SSRPNMAYVLPFMHGKPNRSPPTPSHEAKQFTNEFRAPELKFOYSVKYKAEDLMT 247
 Db 97 IS-SYKNMYLP-FAYSFNSLGVNNKSEA-----KFOLSYKRLLENLGL 141
 QY 248 DSDLMFYTQDSHWQITNGKNSRFRVHDYOPETFLQOPY-SLPLPDGKRYMIGMAVH 306
 Db 142 DEKYIYVYQTSWQIY--EHSSPFRETNYOPEFIDLPLTKDYEFENMLR--VGILH 196
 QY 307 HSNESAK--LSRSMNAYVLMAGMEKMLNLPRIKGRIFREGSGSQPDNDLIDYGY 364
 Db 197 ESNKSGDENTQSNMNIYVSTALINKFLFVPLMYRT--PENKDDDPALHLYMGN 253
 QY 365 GDVRFYOLENKSNIQSVRYNPR--SGKGLQIDVYVPL-GKISGYFOFGYQSGLI 421
 Db 254 FDVVLAV-LDDDFYNLMLNNKLFHNNKGLQVDLGYDIFNNGIYVLYOYFNGYGESLI 312
 QY 422 DYNHEATSFVGLMLN 437
 Db 313 DYNKHLQRLSTGLIS 328

RESULT 5
 Q9SID7 PRELIMINARY; PRT; 292 AA.
 AC Q9SID7;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE PHOSPHOLIPASE A.
 GN PLD4.
 OS Yersinia pseudotuberculosis.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Yersinia.
 OX NCBI_TaxID=633;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YPIII PIB1.
 RA Kariyasev A.V., Minzeler E.A., Williams K.J., Oyston P.C.,
 RA Tibball R.W., Wren B.W.;

RT "Biochip-based Signature-Tagged Mutagenesis: identification and
 RT characterisation of Y. pseudotuberculosis gene pld4 essential for
 RT virulence in mice.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ245393; CAB51586.1;
 DR InterPro: IPR003187; PLAI: 1.
 DR Pfam: PF02253; PLAI: 1.
 DR PRINTS: PR01486; PHPLIPASEA1.
 SQ SEQUENCE 292 AA; 33758 MW; BE712D908ACB6BA5 CRC64;

Query Match 15.3%; Score 360.5; DB 2; Length 292;
 Best Local Similarity 39.2%; Pred. No. 2.3e-23;
 Matches 83; Conservative 34; Mismatches 86; Indels 9; Gaps 5;

QY 229 ELKPOVSKYKAAEDLMTGSDLMFGYTQDSHWQIFNGKNSRFRVHDYOPETFLQPY- 287
 Db 86 EVKFOLSLAFPIWGRIGADNSLGCASYTQSRMQASSESSPRETNYEPQLFLAWSTD 145
 QY 288 YSDLPMDGKVMIGMAVHHNSGESAKLSRSMNAYVLMAGMEKMLNLPRIKGRIFREG 347
 Db 146 YELAGW--TFREVEFGFNHOSNGKADPTSRSMNRYTRVMAQRGNLEIDLPWYRIEED 203
 QY 348 GSGQPDNDLIDYGYGDVRFELYLQ--ENKSNISGTVRYNPRSGKGLQIDVYVPLKGI 406
 Db 204 S---KDNPDITKIMGYRLKLVGALGDSVSLDG--RYNNMTGCGAEMGSIPTIKHV 258
 QY 407 SGYFOIFGYGQSLIDYNHEATSFVGLMLND 438
 Db 259 RFTQVFSGYGESMIDYFNROTGVGIMLND 290

RESULT 6
 Q9L6N9 PRELIMINARY; PRT; 289 AA.
 AC Q9L6N9;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE PLD4 PROTEIN.
 GN PLD4.
 OS Salmonella typhimurium LT2.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=99287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SGSC1412;
 RA Washu;
 RT "The Salmonella typhimurium Genome Sequencing Project.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SGSC1412;
 RA Waterston R.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF233324; AAF3435.1;
 DR InterPro: IPR003187; PLAI: 1.
 DR Pfam: PF02253; PLAI: 1.
 SQ SEQUENCE 289 AA; 32967 MW; DA97F5E1651C49C6 CRC64;

Query Match 15.2%; Score 360; DB 2; Length 289;
 Best Local Similarity 39.7%; Pred. No. 2.6e-23;
 Matches 85; Conservative 29; Mismatches 90; Indels 10; Gaps 4;

QY 226 RAEPLKPOVSKYKAAEDLMTGSDLMFGYTQDSHWQIFNGKNSRFRVHDYOPETFLQ 285
 Db 83 RKDEYKFOLSLAFPIWGRIGADNSLGCASYTQSRMQASSESSPRETNYEPQLFLG 142
 QY 286 PV-YSDLPMDGKVMIGMAVHHNSGESAKLSRSMNAYVLMAGMEKMLNLPRIKGRIF 344
 Db 143 ATDYRFAGW--TLRDVEMGYHDSNGRSDPTSRSMNRLYTRLMADENGMLVLEVAPWYVI- 199

RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
 RA Venter J.C.;
 RT "The complete genome sequence of the gastric pathogen *Helicobacter*
 RT *pylori*.";
 RL Nature 368:539-547(1997).
 DR EMBL: AE000564; AAD07564.1; -
 DR TIGR: HP0499; -
 DR InterPro: IPR003187; PLA1.
 DR Pfam: PF02253; PLA1; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 355 AA; 42486 MW; 461960F15E65AB0A CRC64;

Query Match 10.4%; Score 246.5; DB 2; Length 355;
 Best Local Similarity 25.1%; Pred. No. 2.7e-13;
 Matches 82; Conservative 41; Mismatches 117; Indels 87; Gaps 11;

QY 157 KDAKOLEYAKOFTPLISFDLRNNTPLMSSRPHNPMVLPFIMHGKPNRSPNPTSHE 216
 DB 69 KKIYLNMDYLGTYFLPRYHSF-----TPIFQWYHPNINP----- 102
 QY 217 AKOFTNEFRAPELKFOVSVKVAEDLMCTSDLMFGYTQOSHMOIFNGKNSRPFVHD 276
 DB 103 ---YORNEF-----KFOISFRVPVFRHILMTKGLTLYLATYQTMFQIYNDOQAPMRMIN 154
 QY 277 YQPEIFLQPYVSDLPWDGKV---RMIGCAVHNSG-ESAKLSRSNRAVYLMAGMEKN 332
 DB 155 FMPFLIYVYPI-NKPFQGGKIGNSEIWMQHSNGVGAQCYQPFNK---EGNPENQ 209
 QY 333 LTVMPRI-----WRIKKEGSGSQP-----DNDPD 357
 DB 210 FPGQPIYKDYNGQKDYRMGCRSVSAGQVPRVFLWEKGLKIMAVWYVPYDQSNPN 269
 QY 358 ILDYGEGDYRFLY-----OLENKSINISGTVRYNPRSGKALQLDYVYPLKGIGSYE 410
 DB 270 LIDYMGYGNKIDYRGRHHFELQYDIFQYWRD--RWHGAFRLGYTRINPFV 327
 QY 411 QIFQIGYQSILIDYNHEATSFVGLMLN 437
 DB 328 QWFNGYGDGLYEYDFSNRIGVGIRLN 354

RESULT 10
 Q9ZLX5 PRELIMINARY; PRT; 355 AA.
 AC Q9ZLX5;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE PUTATIVE PHOSPHOLIPASE A1.
 GN PLDA OR JHP0451.
 OS *Helicobacter pylori* J99 (Campylobacter *pylori* J99).
 OC Bacteria; Proteobacteria; epsilon subdivision; *Helicobacter* group;
 OC *Helicobacter*.
 OC NCBI_Taxid=85963;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99120557; PubMed=9923682;
 RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
 RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.W., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RT gastric pathogen *Helicobacter pylori*.";
 RL Nature 397:176-180(1999).
 DR EMBL: AE001479; AAD06029.1; -
 DR InterPro: IPR003187; PLA1.
 DR Pfam: PF02253; PLA1; 1.
 KW Complete proteome.
 SQ SEQUENCE 355 AA; 42329 MW; B3CE9810EBA2FDC9 CRC64;

Query Match 10.2%; Score 240.5; DB 2; Length 355;
 Best Local Similarity 25.1%; Pred. No. 8.9e-13;
 Matches 83; Conservative 44; Mismatches 109; Indels 95; Gaps 13;

QY 157 KDAKOLEYAKOFTPLISFDLRNNTPLMSSRPHNPMVLPFIMHGKPNRSPNPTSHE 216
 DB 69 KKIYLNMDYLGTYFLPRYHSF-----TPIFQWYHPNINP----- 102
 QY 217 AKOFTNEFRAPELKFOVSVKVAEDLMCTSDLMFGYTQOSHMOIFNGKNSRPFVHD 276
 DB 103 ---YORNEF-----KFOISFRVPVFRHILMTKGLTLYLATYQTMFQIYNDOQAPMRMIN 154
 QY 277 YQPEIFLQPYVSDLPWDGKV---RMIGCAVHNSG-ESAKLSRSNR----- 321
 DB 155 FMPFLIYVYPI-NKPFQGGKIGNSEIWMQHSNGVGAQCYQPFNKEGNPENQPPQ 213
 QY 322 -----AYLMAG-----MEWK-NLTVMPRIKGRIFKEGSGSQPDD- 354
 DB 214 PVIYKDYNGQKDYRMGCRSVSAGNALCFVLEKGLKIMAVWYVPY-----PYDQ 265
 QY 355 -NPDILYGYGDYRFLY-----OLENKSINISGTVRYNPRSGKALQLDYVYPLKGIG 406
 DB 266 SNPQLIDYMGYGNKIDYRGRHHFELQYDIFQYWRD--RWHGAFRLGYTRINPFV 323
 QY 407 SGYFOIFQYQSILIDYNHEATSFVGLMLN 437
 DB 324 GIYQWFMNGYGDGLYEYDFSNRIGVGIRLN 354

RESULT 11
 Q9XB53 PRELIMINARY; PRT; 278 AA.
 AC Q9XB53;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE KDU1.
 GN *Erwinia carotovora*.
 OS Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC *Pectobacterium*.
 OC NCBI_Taxid=554;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 39048, GS101;
 RX MEDLINE=98065591; PubMed=9402024;
 RA McGowan S.J., Sebaihia M., O'Leary S., Hardie K.R., Williams P.,
 RA Stewart G.S., Bycroft B.W., Salmond G.P.;
 RT "Analysis of the carboxenem gene cluster of *Erwinia carotovora*:
 RT definition of the antibiotic biosynthetic genes and evidence for a
 RT novel beta-lactam resistance mechanism.";
 RL Mol. Microbiol. 26:545-556(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 39048, GS101;
 RX MEDLINE=98276484; PubMed=9614345;
 RA McGowan S.J., Bycroft B.W., Salmond G.P.;
 RT "Bacterial production of carboxenems and clavams: evolution of beta-
 RT lactam antibiotic pathways.";
 RL Trends Microbiol. 6:203-208(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 39048, GS101;
 RA McGowan S.J.;
 RT Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U17224; AAD38237.1; -
 SQ SEQUENCE 278 AA; 31341 MW; 63769F4F3550E1B7 CRC64;

Query Match 5.2%; Score 123; DB 2; Length 278;
 Best Local Similarity 21.6%; Pred. No. 0.011;
 Matches 74; Conservative 52; Mismatches 128; Indels 88; Gaps 16;

```

QY 51 IDVOSATOSASTDTANPLDEHEPELTYTALLENKMLINCASALNODIMRLACYDTLVHGFT 110
   111 PAVIKTRKSIRLDETIMO-----TIKGPVOVYQETTDPIFLMNEKGM 154
   49 -GIMVDEGLTFDDGIGQFQGVNYFLERREGLINIGGPAKIVDGT--*YEVGNEEL 104
   155 LTKKDAKOLEYAAKQFTPLSLFDLRNNTPLMSSRPHPMYVLPIFMHGKPNRSPNTPS 214
   105 YVKGAKAKALAPS-----SLDSAKPAKLYNSAPAHAVPTPIITODDAIKAPLGCV 155
   215 HEAKQFTNEFRAPPLKQVSVYKAAEDLMGTSDLMFGYT--*QOSHMOIFNGKNSRP 271
   156 KTCNKRTICKLYLPEV-----VETCQ-----LSMGLTRLAEGSNW-----NSMP 194
   272 FRVHDYQPEIFLTQPVYSDLPMDGKVRMIGKGAVHNS-----NGESAKISRSNRAVYLM 326
   195 THTHRRMEVYF---YEDMAEDTIIIFHM-MGEPIETRLVYMHNEQAVISPSMS---IHT 246
   327 GMEWKNLTVMPRIMGRIFKESGSGQPDNDPILDYGYGVDVR 368
   247 GVGTKYNVAF---IMGMT---GENLTFDD---MDHIAMLDLR 278

Db 247 GVGTKYNVAF---IMGMT---GENLTFDD---MDHIAMLDLR 278

RESULT 12
000443 PRELIMINARY: PRT: 1686 AA.
AC 000443:
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PHOSPHOINOSITIDE 3-KINASE.
OS Homo sapiens (Human).
OC Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97479209; PubMed=9337861;
RA Dohn J., Pages F., Vollina S., Rittenhouse S.E., Zvelebil M.J.,
RA Stein R.C., Waterfield M.D.;
RT "Cloning of a human phosphoinositide 3-kinase with a C2 domain that
RT displays reduced sensitivity to the inhibitor wortmannin."
RL Blochem. J. 326:139-147(1997).
DR EMBL: Y13367; CAA73797.1; -.
DR HSSP: P21707; IRYN.
DR InterPro: IPR000008; C2.
DR InterPro: IPR001263; PI3Ka.
DR InterPro: IPR002420; PI3K_C2.
DR InterPro: IPR00341; PI3K_ras_bind.
DR InterPro: IPR000403; PI3_P14_kinase.
DR InterPro: IPR001683; PX.
DR Pfam: PF00168; C2; 1.
DR Pfam: PF00613; PI3Ka; 1.
DR Pfam: PF00792; PI3K_C2; 1.
DR Pfam: PF00794; PI3K_rbd; 1.
DR Pfam: PF00454; PI3_P14_kinase; 1.
DR Pfam: PF00787; PX; 1.
DR SMART: SM00239; C2; 2.
DR SMART: SM00145; PI3Ka; 1.
DR SMART: SM00146; PI3K; 1.
DR SMART: SM00142; PI3K_C2; 1.
DR SMART: SM00144; PI3K_rbd; 1.
DR SMART: SM00312; PX; 1.
DR PROSITE: PS50004; C2_DOMAIN_2; 1.
DR PROSITE: PS00915; PI3_4_KINASE_1; 1.
DR PROSITE: PS00916; PI3_4_KINASE_2; UNKNOWN_1.
DR PROSITE: PS50290; PI3_4_KINASE_3; 1.
SEQUENCE 1686 AA; 150736 MW; E931LC803025C96F CRC64;

```

```

Query Match 5.0%; Score 117; DB 4; Length 1686;
Best Local Similarity 19.2%; Pred. No. 0.48;
Matches 86; Conservative 67; Mismatches 154; Indels 140; Gaps 21;

QY 45 ODNEPLIDV-----OSATOSASTDTANPLDE-----HEPELTYTALLENKMLIN----- 88
   522 EDDETFPVDLANKLVLQIEKCKEAMTRHHVEELLDSTHNOVELALQIDENHRVADQYKAV 581
   89 ---CSALNODIMRLACYDTLVHGFTPAVVKTRKSIRLDETITWQTIKGPVOVYQETTDPI 145
   582 RKICSAID-GVETLAL-----TESVKKLRVAVLPSS-----KTADVT 618
   146 FLMGNEKGLTKKDAKOLEYAAKQFTPLSLFDLRNNTPLMSSRPHPMYV-----LP 199
   619 SLFGGE-----DTSRST-RSLNPNENVOVINOLTTAA 651
   200 IF-----MHGKPNRSPNTPSHEAKQFTPNFRAPLKEFOVSVYKAAEDL---NGTSDLM 252
   652 IYDLRLHANSGRSPDCAOSSKSVKEAMTTTQLOFTIF---AAHGISSMWSVTEKY 707
   253 FGVTQOSHMOIFNGKN-SRPF---RVHDYQPEIFLTQ-----PV-YSDLPMDGKVRMI 300
   708 YLCLSLSH-----NGKDLFRPIQSKKVGTYKNFYLKIMBELIIFPIQISQLPLESVLHLT 763
   301 GMGAVHNSGESAKLSRSN-----RAYLMAGMEWKNLTVMPRIMGRIFK 345
   764 LFGILNOSGSSPDSSNKQKRGPEALGKVLPLCDFRRFLTCG---TKLLYLW----- 812
   346 EGSQSPDNDPILDYGYGVDVRFLYQLENKSNISTVYVYNRSGCALQDLYVPLGK 405
   813 --TSSFTNSVPCTVTKKGYMERIVLQVPSPAFDIITTPQVDRSITIOHNLLETEND 870
   406 ISGYEQIFQYQSLIDYNHEATSFQV 432
   871 IKG-----KLLDILHKDSSUGL 887

Db 871 IKG-----KLLDILHKDSSUGL 887

RESULT 13
063485 PRELIMINARY: PRT: 602 AA.
AC 063485:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE C-RAF ACTIVATED ONCOGENE FUSION PROTEIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87172791; PubMed=3550433;
RA Ishikawa F., Takaku F., Nagao M., Sugimura T.;
RT "Rat c-rac oncogene activation by a rearrangement that produces a
RT fused protein."
RL Mol. Cell. Biol. 7:1226-1232(1987).
CC -1- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: M15428; AAA42002.1; -.
DR InterPro: IPR000719; Euk_kinase.
DR InterPro: IPR002290; Ser_thr_kin_actsite.
DR Pfam: PF00069; pkinase; 1.
DR SMART: SM00221; STYK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; kinase; Oncogene; Serine/threonine-protein kinase;
KW Transmembrane.
SEQUENCE 602 AA; 69209 MW; 88B7BFA90FEB02AC CRC64;

```

Query Match 4.7%; Score 111.5; DB 11; Length 602;
 Best Local Similarity 19.3%; Pred. No. 0.33;
 Matches 93; Conservative 71; Mismatches 174; Indels 145; Gaps 23;


```

Db 85 GLQWPEPEILNDNAYKALSNMDSNMIRLAMY---VGENGYAFNPBLIKQRYIDGIELAI 140
QY 121 RLDETI---WQT-----IKGKPOVYQOETTD-----IF 146
Db 141 ENDWYVYIDWYHVAHAGDPRDPRVYAGAKDFREIALYFNNPHIITYELANEPSSNNNGAG 200
QY 147 LMGNEKGLTKRDAKQLEYAAKQFTPLSLSPDLDRN---NTPLWSSRP-----H 192
Db 201 IPNNEGKWKAVK---EYADPIYEMLRKSGNADDNIIIVGSPNWSORPDLADNPIDDH 255
QY 193 NPMYVLPYFMHGKPNRSPNTPSHKQFTPNFEPFAPELK-----FQVSYKKAEDLMGT 247
Db 256 HTMYTVHFYTGSHAASTESPSE---TPNSEKGNVMSNTRYALENGVAVPATE--WGT 308
QY 248 -----DSDLWFGYTQO-----SHMOIFNGKNS-----RPFYV----- 274
Db 309 SQASGDCGPYFDEADYVLEFLENENNISMANNSLTN--KNEVSGAFTPELGKSNATNLDPG 367
QY 275 --HDYQP-EIFLT-----QPV---YSDLPW--DGKVRMIGMGAVHHSNGE 311
Db 368 PDHVAPEELSLSGEYVRARIKGVNYPEIDRTKYTKVLMDFNDDGTQK--GFGVNSDSPNK 425
QY 312 SAKLSRSNBRAYLMAGME-----WKNLTVMPRIMGR 342
Db 426 ELIAYDNNNTLKVSGLDVSNDSVSDGNFWMANARLSANGWGK 466

```

Search completed: November 30, 2001, 14:26:30
 Job time: 567 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:46:25 ; Search time 72.68 Seconds
(Without alignments)
450.473 Million cell updates/sec

Title: US-09-787-083-2

Perfect score: 442
Sequence: 1 MKVSLSTLTLSTLSCFALLA.....YNHKATPGVGLMNDMMGL 442

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 522463 seqs, 74073290 residues

Word size : 0
Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_1101.*

- 1: /SIDS8/gcgdata/geneseq/AA1980.DAT.*
- 2: /SIDS8/gcgdata/geneseq/AA1981.DAT.*
- 3: /SIDS8/gcgdata/geneseq/AA1982.DAT.*
- 4: /SIDS8/gcgdata/geneseq/AA1983.DAT.*
- 5: /SIDS8/gcgdata/geneseq/AA1984.DAT.*
- 6: /SIDS8/gcgdata/geneseq/AA1985.DAT.*
- 7: /SIDS8/gcgdata/geneseq/AA1986.DAT.*
- 8: /SIDS8/gcgdata/geneseq/AA1987.DAT.*
- 9: /SIDS8/gcgdata/geneseq/AA1988.DAT.*
- 10: /SIDS8/gcgdata/geneseq/AA1989.DAT.*
- 11: /SIDS8/gcgdata/geneseq/AA1990.DAT.*
- 12: /SIDS8/gcgdata/geneseq/AA1991.DAT.*
- 13: /SIDS8/gcgdata/geneseq/AA1992.DAT.*
- 14: /SIDS8/gcgdata/geneseq/AA1993.DAT.*
- 15: /SIDS8/gcgdata/geneseq/AA1994.DAT.*
- 16: /SIDS8/gcgdata/geneseq/AA1995.DAT.*
- 17: /SIDS8/gcgdata/geneseq/AA1996.DAT.*
- 18: /SIDS8/gcgdata/geneseq/AA1997.DAT.*
- 19: /SIDS8/gcgdata/geneseq/AA1998.DAT.*
- 20: /SIDS8/gcgdata/geneseq/AA1999.DAT.*
- 21: /SIDS8/gcgdata/geneseq/AA2000.DAT.*
- 22: /SIDS8/gcgdata/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	442	100.0	442	21	BASB034 amino acid
2	327	74.0	442	21	AAV85268
3	224	50.7	442	21	AAV85269
4	224	50.7	442	21	AAV85270
5	9	2.0	370	21	AAV75156
6	9	2.0	370	21	AAV75157
7	9	2.0	370	21	AAV75158
8	9	2.0	374	21	AAV70629
9	9	2.0	375	21	AAV70628
10	8	1.8	278	21	AAV1137
11	8	1.8	283	21	AAV1136

12	8	1.8	299	21	AAV1135	Arabidopsis thalia
13	7	1.6	16	15	AAV52127	Mouse light chain
14	7	1.6	33	22	AAV76059	Human colon cancer
15	7	1.6	50	21	AAV79483	Rat alpha tubulin
16	7	1.6	64	21	AAV60322	Arabidopsis thalia
17	7	1.6	82	21	AAV01488	Human secreted pro
18	7	1.6	106	21	AAV07608	Arabidopsis thalia
19	7	1.6	105	21	AAV51952	Arabidopsis thalia
20	7	1.6	108	21	AAV37811	Arabidopsis thalia
21	7	1.6	116	21	AAV87313	Human signal pepti
22	7	1.6	136	21	AAV37810	Arabidopsis thalia
23	7	1.6	157	21	AAV29867	Arabidopsis thalia
24	7	1.6	158	21	AAV79176	Haematopoietic ste
25	7	1.6	169	20	AAV34661	Chlamydia pneumoni
26	7	1.6	181	21	AAV37809	Arabidopsis thalia
27	7	1.6	186	21	AAV07607	Arabidopsis thalia
28	7	1.6	186	21	AAV51951	Arabidopsis thalia
29	7	1.6	190	21	AAV07606	Arabidopsis thalia
30	7	1.6	190	21	AAV07606	Arabidopsis thalia
31	7	1.6	268	13	AAV21542	Envelope protein o
32	7	1.6	302	13	AAV24254	Protease inhibitor
33	7	1.6	316	21	AAV39500	Arabidopsis thalia
34	7	1.6	324	21	AAV39499	Arabidopsis thalia
35	7	1.6	326	13	AAV24356	Protease inhibitor
36	7	1.6	365	21	AAV39498	Arabidopsis thalia
37	7	1.6	381	20	AAV60241	Human endometrium
38	7	1.6	448	21	AAV91960	Human cytoskeleton
39	7	1.6	523	21	AAV44285	Arabidopsis thalia
40	7	1.6	552	16	AAV67007	Firefly luciferase
41	7	1.6	666	20	AAV21871	Amino acid sequenc
42	7	1.6	683	20	AAV21869	Amino acid sequenc
43	7	1.6	684	20	AAV21873	Amino acid sequenc
44	7	1.6	684	20	AAV21875	Amino acid sequenc
45	7	1.6	715	20	AAV21860	Amino acid sequenc

ALIGNMENTS

RESULT 1	AAV85268	standard; Protein: 442 AA.
AC	AAV85268;	
XX		
XX		
DT	29-JUN-2000	(first entry)
XX		
DE	BASB034 amino acid sequence #1.	
XX		
KW	Moraxella catarrhalis infection; BASB034: diagnosis; staging;	
KW	vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;	
KW	sinusitis; nosocomial infection; invasive disease; chronic otitis media;	
KW	hearing loss; antibacterial drug.	
OS	Moraxella catarrhalis.	
XX		
PN	W0200015802-A1.	
PD	23-MAR-2000.	
XX		
PF	14-SEP-1999;	99WO-EP06781.
XX		
PR	14-SEP-1998;	98GB-0020002.
PA	(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.	
PI	Ruelle J;	
XX		
DR	WPI: 2000-271440/23.	
XX	N-PSDB: AAA10700.	
PT	Novel BASB034 polynucleotides and polypeptides from Moraxella	
PT	catarrhalis used to prepare vaccines against bacterial infections	

XX Claim 3; Fig 2; 106pp; English.
 PS
 CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 CC strain Mc2931 (ATCC 43617). The invention relates to BASB034 polypeptides
 CC from M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2965. The
 CC BASB034 polynucleotides and polypeptides may be employed as research
 CC reagents and material for the discovery of treatments and diagnostics for
 CC diseases, particularly human diseases. They are particularly used to
 CC diagnose and treat M. catarrhalis infections. They can be used for
 CC diagnosis of disease, staging of disease, or determining response of an
 CC infectious organism to drugs. The polynucleotides may be used as a source
 CC for hybridization probes, and for screening of genetic mutations.
 CC serotype, organism or strain identification, identification of mutations
 CC in BASB034 sequences, and as components of arrays which are useful for
 CC diagnostic and prognostic purposes. The polypeptides can be used to
 CC produce antibodies. The polypeptides can also be used in vaccine
 CC formulations, and to identify agonists and antagonists. The polypeptides,
 CC antibodies, agonists and antagonists (which are bacteriostatic) are used
 CC for the treatment and prevention of diseases such as otitis media in
 CC infants and children, pneumonia in elderlies, sinusitis, nosocomial
 CC infections and invasive diseases, and chronic otitis media with hearing
 CC loss. The polypeptides, agonists and antagonists are also used for
 CC screening of antibacterial drugs. The BASB034 products of the invention
 CC can be used screen for new antibacterial compounds that may target
 CC resistant bacteria.
 CC
 XX
 SQ Sequence 442 AA;

Query Match 100.0%; Score 442; DB 21; Length 442;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSLTILSLISCAIILAIQQAQAVNPVAFVDEVRSENDIGQDNELPIDVQATQSA 60
 DB 1 mkvsltilsliscailaiaiqgaqavnpvatfdevrsendigqdnelpidvqatqsa 60
 QY 61 STDTANPLDEHEBELYTTALENKTMKLNCSALNODIMRLACYDTLVHGETPAVIKTRKST 120
 DB 61 stdtanpldehebellyttalenktmlncsalnqdimrlacydltlvhgetpaviktrkrsi 120
 QY 121 RLDETWMQIKGKPOYIYETTDPIFLMGNEKMLTKKAKOLEYAAKOTPTLSFDD 180
 DB 121 rldetwmqikgkpoiyiyettdpiflmgnekmltkkakokeyaakotptlsfdd 180
 QY 121 rldetlwtqtkgkpvlygettdpiflmgnekymtkkakeleyaakqftpsisfdd 180
 DB 181 RNNTPLMSSRPNNPNTVLPILFMHGKPNRSPNTPSHBAKQFTPEFRAPELKEFVSKVKA 240
 DB 181 rnntpplmssrpnnpntvlpilfmhgkpnrsnptshbakqftpefrapelkftvskvka 240
 QY 241 AEDLMCTSDLMFGYQOQSHWQJFNCKNSRPFVNDYDPEIFLTQPVYSDLPMDGKVRMT 300
 DB 241 aedlmctsdlmfgyqoqshwqjfncknsrpfvndydpelfltqpvysdldpmdgkvrm 300
 QY 241 aedltwtdadlrfytcqgshwqjfncknsrpfvndydpelfltqpvysdldpmdgkvrm 300
 DB 301 GMGAVHNSGSEKSLRSNNRAYLTMAGMEKNLTVMPTMGRTFRKSGSGSPDNDPDIID 360
 DB 301 gmgavhnsgekslrsnnraylimgmeknltvmptmgrtfrksgsgspdpndpdiid 360
 QY 361 YVGYGVNRELXYOLENKSNTSGTVRNPYRSGKALQLDYVYPLGKSGFQIOTQGVQST 420
 DB 361 yvgygvnrelxyolenksntsgtvrnpysrgkalgldyvypylgksgfqiotsqvgst 420
 QY 421 IDYNHEATSEFGVGLMDNMWGL 442
 DB 421 idynheatsfgyglmdnmwgl 442

RESULT 2
 AAY85271
 ID AAY85271 standard; Protein; 442 AA.
 XX
 AC AAY85271;
 XX

DT 29-JUN-2000 (first entry)
 XX
 DE BASB034 amino acid sequence #4.
 XX
 KW Moraxella catarrhalis infection; BASB034; diagnosis; staging;
 KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KW hearing loss; antibacterial drug.
 XX
 OS Moraxella catarrhalis.
 PN WO200015802-A1.
 XX
 XX 23-MAR-2000.
 PD
 PD 14-SEP-1999; 99WO-EP06781.
 PF
 PF 14-SEP-1998; 98GB-0020002.
 PR
 PR (SMK) SMITHKLINE BEECHAM BIOLOGICALS.
 PA
 PA Ruelle J;
 PI
 PI WPI: 2000-271440/23.
 DR
 DR N-PSDB; AAA10703.
 XX
 PT Novel BASB034 polynucleotides and polypeptides from Moraxella
 PT catarrhalis used to prepare vaccines against bacterial infections
 XX
 PS Claim 3; Page 69; 106pp; English.

QY 1 MKVSLTILSLISCAIILAIQQAQAVNPVAFVDEVRSENDIGQDNELPIDVQATQSA 60
 DB 1 mkvsltilsliscailaiaiqgaqavnpvatfdevrsendigqdnelpidvqatqsa 60
 QY 61 STDTANPLDEHEBELYTTALENKTMKLNCSALNODIMRLACYDTLVHGETPAVIKTRKST 120
 DB 61 stdtanpldehebellyttalenktmlncsalnqdimrlacydltlvhgetpaviktrkrsi 120
 QY 121 RLDETWMQIKGKPOYIYETTDPIFLMGNEKMLTKKAKOLEYAAKOTPTLSFDD 180
 DB 121 rldetwmqikgkpoiyiyettdpiflmgnekmltkkakokeyaakotptlsfdd 180
 QY 121 rldetlwtqtkgkpvlygettdpiflmgnekymtkkakeleyaakqftpsisfdd 180
 DB 181 RNNTPLMSSRPNNPNTVLPILFMHGKPNRSPNTPSHBAKQFTPEFRAPELKEFVSKVKA 240
 DB 181 rnntpplmssrpnnpntvlpilfmhgkpnrsnptshbakqftpefrapelkftvskvka 240
 QY 241 AEDLMCTSDLMFGYQOQSHWQJFNCKNSRPFVNDYDPEIFLTQPVYSDLPMDGKVRMT 300
 DB 241 aedlmctsdlmfgyqoqshwqjfncknsrpfvndydpelfltqpvysdldpmdgkvrm 300
 QY 241 aedltwtdadlrfytcqgshwqjfncknsrpfvndydpelfltqpvysdldpmdgkvrm 300
 DB 301 GMGAVHNSGSEKSLRSNNRAYLTMAGMEKNLTVMPTMGRTFRKSGSGSPDNDPDIID 360
 DB 301 gmgavhnsgekslrsnnraylimgmeknltvmptmgrtfrksgsgspdpndpdiid 360
 QY 361 YVGYGVNRELXYOLENKSNTSGTVRNPYRSGKALQLDYVYPLGKSGFQIOTQGVQST 420
 DB 361 yvgygvnrelxyolenksntsgtvrnpysrgkalgldyvypylgksgfqiotsqvgst 420
 QY 421 IDYNHEATSEFGVGLMDNMWGL 442
 DB 421 idynheatsfgyglmdnmwgl 442

Query Match 74.0%; Score 327; DB 21; Length 442;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 427; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 15 CFAIIIAIQQAQAVNPVAFVDEVRSENDIGQDNELPIDVQATQSASTDTPANPLDEHEPE 74
 DB 15 cfaillaiaqgaqavnpvatfdevrsendigqdnelpidvqatqsastdtanpldehepe 74
 QY 75 LYTTALENKTMLINCALNODIMRLACYDTLVHGETPAVIKTRKSTRLDETWMQIKGK 134
 DB 75 lyttalenktmlincsalnqdimrlacydltlvhgetpaviktrksrldetlwtqtkgk 134
 QY 135 OVIYQETTDPIFLMGNEKMLTKKAKOLEYAAKOTPTLSFDDRLNNTPLMSSRPNNP 194
 DB 135 oviyqettdpiflmgnekmltkkakokeyaakotptlsfddrlnntplmssrpnnp 194

OY 195 MYVLPIFMHGKPNRSPNTPSHEAKOFTPNFRAPBLKFOVSKYKAAEDIMGTDSDLMFG 254
 DB 195 MYVLPILFMHGKPNRSPNTPSHEAKQFIPNEITRAPELKITQVSKVKAEDLWGTDSDLWFG 254
 OY 255 YTOQSHMOIFNGKNSRPRVHDYOPEIFLTQPVYSDLPWDGKVRMIGMAGVHNSNGESAK 314
 DB 255 YTGQSHWQIFNGKNSRPRVHDYOPEIFLTQPVYSDLPWDGKVRMIGMAGVHNSNGESAK 314
 OY 315 LSRSMNRAVLMAGMEKMLTYVPRIMGRIFKESGSGQPDNDLIDYGYGQDREFLYOLE 374
 DB 315 LSRSMNRAVLMAGMEKMLTYVPRIMGRIFKESGSGQPDNDLIDYGYGQDREFLYOLE 374
 OY 375 NKSNTSGVRNPNRSGKALOLDYVPRIGKISGTFQIFQCGQSLIDYNEHATSFQVGL 434
 DB 375 NKSNTSGVRNPNRSGKALOLDYVPRIGKISGTFQIFQCGQSLIDYNEHATSFQVGL 434
 OY 435 MLNDMNGL 442
 DB 435 mlndmngl 442
 RESULT 3
 ID AAY85269 standard; Protein: 442 AA.
 AC AAY85269;
 DT 29-JUN-2000 (first entry)
 DE BASB034 amino acid sequence #2.
 KW Moraxella catarrhalis infection; BASB034; diagnosis; staging;
 KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KW hearing loss; antibacterial drug.
 OS Moraxella catarrhalis.
 PN WO200015802-A1.
 PD 23-MAR-2000.
 PE 14-SEP-1999; 99WO-EP06781.
 PR 14-SEP-1998; 98GB-0020002.
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 PI Ruelle J;
 DR WPI: 2000-271440/23.
 DR N-PSDB: AAA10701.
 PT Novel BASB034 polynucleotides and polypeptides from Moraxella
 catarrhalis used to prepare vaccines against bacterial infections -
 Claim 3; Page 67; 106pp; English.
 This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 strain Mc2908. The invention relates to BASB034 polypeptides from
 M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
 polynucleotides and polypeptides may be employed as research reagents and
 material for the discovery of treatments and diagnostics for diseases,
 particularly human diseases. They are particularly used to diagnose and
 treat M. catarrhalis infections. They can be used for diagnosis of
 disease, staging of disease, or determining response of an infectious
 organism to drugs. The polynucleotides may be used as a source for
 hybridization probes, and for screening of genetic mutations, serotype,
 organism or strain identification, identification of mutations in BASB034
 sequences, and as components of arrays which are useful for diagnostic
 and prognostic purposes. The polypeptides can be used to produce
 antibodies. The polypeptides can also be used in vaccine formulations,

CC and to identify agonists and antagonists. The polypeptides, antibodies,
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elders, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.
 CC
 XX Sequence 442 AA;
 SQ
 Query Match 50.7%; Score 224; DB 21; Length 442;
 Best Local Similarity 100.0%; Pred. No. 9.2e-209;
 Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 219 QFTPNFRAPBLKFOVSKYKAAEDLMGTDSDLMFGYTOQSHMOIFNGKNSRPRVHDYQ 278
 DB 219 qftpnfrapblkitqvsyvkkaedlwgtdsdlwfgyqgshwqifngknsrprfvhdq 278
 OY 279 PEIFLTQPVYSDLPWDGKVRMIGMAGVHNSNGESAKLSRSMNRAVLMAGMEKMLTYVPR 338
 DB 279 peifltqpvysdlpwdgkvrmlgmagvhnsgesaklsrsmnraylmagmekmltvmpr 338
 OY 339 IMGRIFKESGSGQPDNDLIDYGYGQDREFLYOLENKSNTSGTVRYNPNRSGKALOLDY 398
 DB 339 iwmgrifksgsgqpdndlidygygdvrflyoleksnsgtvyrynpnrsgkalgldy 398
 OY 399 VYPLGKISGYFQIFQCGQSLIDYNEHATSFQVGLMLNDMNGL 442
 DB 399 vyplkgisgyfqifqgygslidynheatsfyglmlndmngl 442
 RESULT 4
 ID AAY85270 standard; Protein: 442 AA.
 AC AAY85270;
 DT 29-JUN-2000 (first entry)
 DE BASB034 amino acid sequence #3.
 KW Moraxella catarrhalis infection; BASB034; diagnosis; staging;
 KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KW hearing loss; antibacterial drug.
 OS Moraxella catarrhalis.
 PN WO200015802-A1.
 PD 23-MAR-2000.
 PE 14-SEP-1999; 99WO-EP06781.
 PR 14-SEP-1998; 98GB-0020002.
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 PI Ruelle J;
 DR WPI: 2000-271440/23.
 DR N-PSDB: AAA10702.
 PT Novel BASB034 polynucleotides and polypeptides from Moraxella
 catarrhalis used to prepare vaccines against bacterial infections -
 Claim 3; Page 68; 106pp; English.
 This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 strain Mc2913. The invention relates to BASB034 polypeptides from

CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
CC polynucleotides and polypeptides may be employed as research reagents and
CC material for the discovery of treatments and diagnostics for diseases,
CC particularly human diseases. They are particularly used to diagnose and
CC treat M. catarrhalis infections. They can be used for diagnosis of
CC disease, staging of disease, or determining response of an infectious
CC organism to drugs. The polynucleotides may be used as a source for
CC hybridization probes, and for screening of genetic mutations, serotype,
CC organism or strain identification, identification of mutations in BASB034
CC sequences, and as components of arrays which are useful for diagnostic
CC and prognostic purposes. The polypeptides can be used to produce
CC antibodies. The polypeptides can also be used in vaccine formulations,
CC and to identify agonists and antagonists. The polypeptides, antibodies,
CC agonists and antagonists (which are bacteriostatic) are used for the
CC treatment and prevention of diseases such as otitis media in infants and
CC children, pneumonia in elderly, sinusitis, nosocomial infections and
CC invasive diseases, and chronic otitis media with hearing loss. The
CC polynucleotides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.

SQ Sequence 442 AA;

Query Match	50.7%	Score 224	DB 21	Length 442
Best Local Similarity	100.0%	Pred. No.	9.2e-209	
Matches 224	Conservative	0	Mismatches	0
			Indels	0
			Gaps	0

Qy	219	QFENNEFAPARLKTQSVKVVAAADLNGTSDLMEGVTQOSHQIENFGKNSRPRRNDYQ	278
Db	219	qftrnefrfapeklqvsykkvnaaealvgltbsdlmfyrlqsgshwllngknsrfrlrvndyg	278
Qy	279	PEIFLTQPVYSDLLPMDGCKVRMIGICGAHNHSGESAKLSRSNNRAYLAAEGEMKKLTWMPR	338
Db	279	peifltqpvysddlpxdvgkvvmlmgagvhhnsgeaaklsrswntaylmaegewknltwmptr	338
Qy	339	IGWRIFKEGSGSQPDNDNDLIDYGYGDVRELYOLENKSNIISGTVRNPSPSGKALOLDY	398
Db	339	lwgrifkegsgsqdpndndlddygygdvrflylgnknsisgvcvrynprrsgkalldy	398
Qy	399	VYPLGKGTSGTFOJFGYGGGSLIDYHNHATSFGYGLMLNDMMKL	442
Db	399	vyp.lgkgtsgyfqj.fgyyggslidynhneatsfyg.lmlndmmgl	442

RESULT 5

AAV75156 standard; Protein; 370 AA

AC AAY75156;

DT 21-MAR-2000 (first entry)

DE *Neisseria gonorrhoeae* ORF 582 protein sequence SEQ ID NO:1786.

Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine; antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia, antibacterial; gene therapy.

OS *Neisseria gonorrhoeae*.

PN W09957280-A2

PD 1.1-NOV-1999

PF 30-APR-1999; 99WO-US09346.

PR 01-MAY-1998; 98US-0083758;
PR 31-JUL-1998; 98US-0094869;
PR 02-SEP-1998; 98US-0094894;
PR 02-SEP-1998; 98US-0099062;
PR 09-OCT-1998; 98US-0103749;

PR	09-OCT-1998;	98US-0103794.
PR	09-OCT-1998;	98US-0103796.
PR	25-FEB-1999;	99US-0121528.

PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.

PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M,
PI Petersen J, Pizzo M, Rappuoli R, Ratti G, Scalato E, Scarselli M,
PI Tettelin H, Venter JC;

DR WPI; 2000-062150/05.
DR N-PSDB; AA253918.

PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics -

PS Claim 2; Page 903; 1453pp; English.

AAZ5301, oAAZ54556, AAZ54577 to AAZ54615, and AAZ74253 to AAZ75941 represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polypeptides and polypeptides. AAZ54537 and AAZ54576 and AAZ54616 to AAZ55473 represent PCR primers used in the exemplification of the present invention. The polypeptides, the polynucleotides, antibodies and compositions of the invention can be used as vaccines, as diagnostic reagents, and as immunogenic compositions. The polypeptides can be used in the manufacture of medicaments for treating or preventing infection due to *Neisseria* bacteria (e.g. meningitis and septicæmia), to detect the presence of *Neisseria* bacteria, or to raise antibodies. They may also be used to screen for agonists or antagonists, which may themselves have use as antibacterial agents. The polynucleotides of the invention may also be used in gene therapy protocols.

SQ Sequence 370 AA;

Query Match	2.0%;	Score 9;	DB 21;	Length 370;
Best Local Similarity	100.0%;	Pred.No. 2.3;		
Matches 9; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

QY	279	PEIFLTQPV	287
Db	209	peifltqpv	217

RESULT 6

AAV75157 standard; Protein; 370 AA.

AC AAY75157;

DT 21-MAR-2000 (first entry)

DE Nelsleria meningitidis ORF 582 protein sequence SEQ ID NO:1788.

KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy.

05 *Neisseria meningitidis*.

PN W09957280-A2

PD 11-NOV-1999

PF 30-APR-1999; 99WO-US09346.

PR 01-MAY-1998; - 98US-0083758;
PR 31-JUL-1998; 98US-0094869;
PR 02-SEP-1998; 98US-0098394;
PR 02-SEP-1998; 98US-0099662;
PR 09-OCT-1998; 98US-0103749;
PR 09-OCT-1998; 98US-0103794;

PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX
 DR WPI; 2000-062150/05.
 DR N-PSDB; AA253919.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics -
 XX
 PS Claim 2; Page 904; 1453pp; English.
 XX
 CC AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
 CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA255473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 XX
 SQ Sequence 370 AA;

Query Match 2.0%; Score 9; DB 21; Length 370;
 Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 PEFLTPGV 287
 |||||||||
 Db 209 pelfltqpv 217

RESULT 7
 AAY75158
 ID AAY75158 standard; Protein; 370 AA.
 XX
 AC AAY75158;
 XX
 DT 21-MAR-2000 (first entry)
 XX
 DE Neisseria meningitidis ORF 582 protein sequence SEQ ID NO:1790.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KM antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
 KM antibacterial; gene therapy.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO9957280-A2.
 PD
 PD 11-NOV-1999.
 XX
 PF 30-APR-1999; 99WO-US09346.
 XX
 PR 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.

PR 25-FEB-1999; 99US-0121528.
 XX (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX
 DR WPI; 2000-062150/05.
 DR N-PSDB; AA253920.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics -
 XX
 PS Claim 2; Page 905; 1453pp; English.
 XX
 CC AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
 CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA255473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 XX
 SQ Sequence 370 AA;

Query Match 2.0%; Score 9; DB 21; Length 370;
 Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 PEFLTPGV 287
 |||||||||
 Db 209 pelfltqpv 217

RESULT 8
 AAY70629
 ID AAY70629 standard; Protein; 374 AA.
 XX
 AC AAY70629;
 XX
 DT 18-JUL-2000 (first entry)
 XX
 DE Neisseria meningitidis serogroup B strain H44/76 BASB033 protein.
 XX
 KW BASB033; diagnosis; prophylaxis; treatment; antibacterial; vaccine;
 KM Neisseria meningitidis infection.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO200015801-A1.
 PD
 PD 23-MAR-2000.
 XX
 PF 09-SEP-1999; 99WO-EP06718.
 XX
 PR 14-SEP-1998; 98GB-0020003.
 XX
 PA (SMIR) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 PI Ruelle J;
 XX
 DR WPI; 2000-271439/23.
 DR N-PSDB; AA252134.

PT Isolated BASB033 polypeptides and polynucleotides of *Neisseria*
PT meningitidis, useful for diagnosis, prophylaxis and treatment of *N.*
PT meningitidis infection -
XX
XX
PS Claim 4: Page 59; 93pp; English.
XX
CC The present sequence is a BASB033 protein from
CC *Neisseria meningitidis* serogroup B strain H44/76. The protein
CC shows homology to the *Klebsiella pneumoniae* outer membrane
CC phospholipase A. The present sequence is useful for diagnosis,
CC prophylaxis and treatment of *N. meningitidis* infection. It may also be
CC used for the discovery and development of antibacterial compounds and
CC in vaccine compositions.
XX
SQ Sequence 374 AA;

Query Match 2.0%; Score 9; DB 21; Length 374;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 PEIFLTQPV 287
DB 213 pefltcpv 221

RESULT 9
AAV70628
ID AAV70628 standard; Protein; 375 AA.
XX
XX AAV70628;
XX
DT 18-JUL-2000 (first entry)
XX
DE *Neisseria meningitidis* serogroup B strain ATCC13090 BASB033 protein.
XX
XX BASB033; diagnosis, prophylaxis; treatment; antibacterial; vaccine;
XX *Neisseria meningitidis* infection.
XX
OS *Neisseria meningitidis*.
XX
PN WO200015801-A1.
XX
PD 23-MAR-2000.
XX
PF 09-SEP-1999; 99WO-EP06718.
XX
PR 14-SEP-1998; 98GB-0020003.
XX
PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruehle J;
XX
DR MPI: 2000-271439/23.
XX
DR N-PSDB; AA252133.
XX
PT Isolated BASB033 polypeptides and polynucleotides of *Neisseria*
PT meningitidis, useful for diagnosis, prophylaxis and treatment of *N.*
PT meningitidis infection -
XX
XX Claim 4: Page 58; 93pp; English.
XX
XX The present sequence is a BASB033 protein from
XX *Neisseria meningitidis* serogroup B strain ATCC13090. The protein
XX shows homology to the *Klebsiella pneumoniae* outer membrane
XX phospholipase A. The present sequence is useful for diagnosis,
XX prophylaxis and treatment of *N. meningitidis* infection. It may also be
XX used for the discovery and development of antibacterial compounds and
XX in vaccine compositions.
SQ Sequence 375 AA;

Query Match 2.0%; Score 9; DB 21; Length 375;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 PEIFLTQPV 287
DB 214 pefltcpv 222

RESULT 10
AAG11137
ID AAG11137 standard; Protein; 278 AA.
XX
XX AAG11137;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 9741.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP103405-A2.
XX
PN 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PF 25-FEB-1999; 99US-0121825.
XX
PR 05-MAR-1999; 99US-0123180.
XX
PR 09-MAR-1999; 99US-0123548.
XX
PR 23-MAR-1999; 99US-0125788.
XX
PR 29-MAR-1999; 99US-0126264.
XX
PR 29-MAR-1999; 99US-0126785.
XX
PR 01-APR-1999; 99US-0127462.
XX
PR 06-APR-1999; 99US-0128234.
XX
PR 08-APR-1999; 99US-0128714.
XX
PR 16-APR-1999; 99US-0129845.
XX
PR 19-APR-1999; 99US-0130077.
XX
PR 21-APR-1999; 99US-0130449.
XX
PR 23-APR-1999; 99US-0130510.
XX
PR 28-APR-1999; 99US-0130891.
XX
PR 30-APR-1999; 99US-0131449.
XX
PR 30-APR-1999; 99US-0132048.
XX
PR 04-MAY-1999; 99US-0132407.
XX
PR 04-MAY-1999; 99US-0132484.
XX
PR 05-MAY-1999; 99US-0132485.
XX
PR 06-MAY-1999; 99US-0132486.
XX
PR 07-MAY-1999; 99US-0132487.
XX
PR 11-MAY-1999; 99US-0132863.
XX
PR 14-MAY-1999; 99US-0134219.
XX
PR 14-MAY-1999; 99US-0134219.
XX
PR 14-MAY-1999; 99US-0134221.
XX
PR 14-MAY-1999; 99US-0134370.
XX
PR 18-MAY-1999; 99US-0134768.
XX
PR 19-MAY-1999; 99US-0134941.
XX
PR 20-MAY-1999; 99US-0135124.
XX
PR 21-MAY-1999; 99US-0135153.
XX
PR 24-MAY-1999; 99US-0135629.
XX
PR 25-MAY-1999; 99US-0136021.
XX
PR 28-MAY-1999; 99US-0136392.
XX
PR 01-JUN-1999; 99US-0137222.
XX
PR 03-JUN-1999; 99US-0137528.
XX
PR 04-JUN-1999; 99US-0137502.
XX
PR 07-JUN-1999; 99US-0137724.
XX
PR 08-JUN-1999; 99US-0138094.
XX
PR 10-JUN-1999; 99US-0138540.
XX
PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 8; DB 21; Length 278;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;
Qy 174 slsfdldr 181
|||||||
Db 271 slsfdldr 278

RESULT 11
AAG1136
ID AAG1136 standard; Protein; 283 AA.
XX AAG1136;
AC AAG1136;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 9740.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148655.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 8; DB 21; Length 283;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 SLSEFIDR 181
| | | | | | | | | |
Db 276 SLSEFIDR 283

RESULT 12
AAG1135
ID AAG1135 standard; Protein; 299 AA.
XX
AC AAG1135;

XX 17-OCT-2000 (first entry)
DT Arabidopsis thaliana protein fragment SEQ ID NO: 9739.
DE
XX Protein identification: signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
PD 06-SEP-2000.
XX
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137702.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139494.
PR 18-JUN-1999; 99US-0139495.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.

PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161350.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 8; DB 21; Length 299;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 SLSFDLDR 181
Db 292 slsfdldr 299

RESULT 13
AARS2127
ID AARS2127 standard; Peptide: 16 AA.
XX
AC AARS2127;
XX
DT 27-SEP-1996 (first entry)
XX
XX Mouse light chain surface patch S03410.
DE
XX
KW antibody; humanised; murine; human; heavy chain; light; variable;
framework region; complementarity determining region; reshaping;

PA	(HOMA-) HUMAN GENOME SCI INC.
PI	Ruben SM, Barash SC, Birse CE, Rosen CA;
XX	WPI: 2001-235357/24.
DR	N-PSDB: AAH35474.
XX	
PT	Nucleic acids encoding 4277 human colon cancer-associated polypeptides, useful for preventing, diagnosing and/or treating colorectal cancers -
PS	Claim 11: Page 8285-8286; 9803bp; English.
XX	
CC	AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon cancer-associated nucleic acid molecules (N) and proteins (P), where the proteins are collectively known as colon cancer antigens. The colon cancer antigens have cytostatic activity and can be used in gene therapy and vaccine production. N and P may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate P expression. For example, N and P may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of P by expressing inactive proteins or to supplement the patient's own production of P. Additionally, N may be used to produce the colon cancer-associated P, by inserting the nucleic acids into a host cell and culturing the cell to express the proteins. N and P can be used in the prevention, diagnosis and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent sequences used in the exemplification of the present invention.
CC	N.B. Pages 666 to 682 and page 7053 of the sequence listing were missing at time of publication, meaning no sequences are present for SEQ ID NO:1027 to 1052, 7921 and 7922.
CC	
XX	Sequence 33 AA:
XX	
XX	Query Match 1.6%; Score 7; DB 22; Length 33;
XX	Best Local Similarity 100.0%; Pred. No. 26;
XX	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	8 LTLSTLS 14 22 ltlstls 28
Db	
RESULT 15	
AA79483	
ID	AA79483 standard; peptide; 50 AA.
XX	
XX	AA79483;
XX	
XX	01-AUG-2000 (first entry)
DE	
XX	Rat alpha tubulin variant peptide (aa314-363).
XX	
KW	Rat: alpha tubulin; mutant; mutein; variant;
KW	metabolic X syndrome; hypertension; stroke; diabetes;
KW	insulin resistance; obesity; dyslipidemia; anorectic;
KW	hypotensive; cerebroprotective; vasotropic; antidiabetic;
XX	antihypemic; gene therapy; spontaneously hypertensive; SHR.
OS	
XX	Rattus sp.
XX	
XX	Key Location/Qualifiers
FT	1..50
FT	Peptide /note= "corresponds to residues 314-363 of full-length protein"
FT	27
FT	Misc-difference /note= "replaces Thr in wild-type"
XX	
XX	WO200018918-A2.
XX	
XX	06-APR-2000.
XX	

PF 28-SEP-1999; 99WO-US22494.
XX
PR 28-SEP-1998; 98US-0161939.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shinkets RA:
XX WPI; 2000-303450/26.
DR
XX
PT Novel genes encoding e.g. CD36, SGLT2, and kynurenine aminotransferase,
PT which are predictive and therapeutic for stroke, hypertension, diabetes
PT and obesity
XX
XX
PS Claim 2; Fig 4G; 79pp; English.
XX
XX This sequence represents amino acids 314-363 of alpha-tubulin
CC from a spontaneously hypertensive rat (SHR). This region of
CC alpha-tubulin carries a Ser for Thr amino acid substitution when
CC compared to the corresponding region in control Wistar Kyoto (WKY)
CC rats (see AAY/9484). The alpha-tubulin gene is differentially
CC expressed in SHR, SHR-stroke prone (SP) and WKY rats. It may be
CC involved in a predisposition to vascular injury. The SHR rat
CC is an animal model for human metabolic X syndrome. Genes encoding
CC sodium dependent glucose cotransporter, kynurenine aminotransferase,
CC CD36, aldolase A, atrial natriuretic factor, alpha-cardiac myosin
CC and alpha-tubulin were identified as being potentially associated
CC with hypertension, obesity and insulin resistance. These
CC proteins, including isolated proteins comprising the present
CC sequence, are used for treating, preventing and diagnosing
CC ischemic and metabolic diseases and disorders, such as stroke,
CC hypertension, diabetes and obesity, especially insulin resistivity,
CC dyslipidemia and ischemic stroke (all claimed). Polynucleotides
CC encoding them may also be used in gene therapy and antisense
CC therapy protocols.
XX
SQ Sequence 50 AA;

Query Match 1.6%; Score 7; DB 21; Length 50;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 114 IKTKRSI 120
IIIIIIII
DB 22 Iktkrsi 28

Search completed: May 9, 2002, 12:54:58
Job time: 513 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:56:32 ; Search time 44.37 Seconds

(without alignments)
758,827 Million cell updates/sec

Title: US-09-787-083-4

Perfect score: 442
Sequence: 1 MKVSLSTLTLSTLPCFAILA.....YVHEATSPGVGLMDNMGL 442

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	2.0	382	2 E81195	phospholipase A1,
2	9	2.0	409	2 H81831	probable phospholip
3	8	1.8	286	2 B36971	outer membrane pho
4	8	1.8	299	2 T47989	RAV-like protein -
5	8	1.8	1396	2 S36851	L-shaped tail fibre
6	7	1.6	122	2 E75377	hypothetical prote
7	7	1.6	155	2 A29659	beta-lactoglobulin
8	7	1.6	158	2 E86498	pts iia protein (l
9	7	1.6	158	2 E72124	tubulin alpha-1 ch
10	7	1.6	197	2 A54506	hypothetical prote
11	7	1.6	219	2 T19438	tubulin alpha chai
12	7	1.6	240	2 A61544	coat protein - pot
13	7	1.6	264	2 A44959	genome polypotein
14	7	1.6	267	2 A60366	capsid protein - p
15	7	1.6	267	2 S14001	coat protein - pot
16	7	1.6	267	2 S26630	tail fiber protein
17	7	1.6	267	2 JCL527	coat protein - pot
18	7	1.6	267	2 S13239	tail fiber protein
19	7	1.6	270	2 S73734	abc transport ATP-
20	7	1.6	273	2 T36237	transaminase B hom
21	7	1.6	288	2 E72576	probable phosphor
22	7	1.6	292	2 T36237	probable ABC trans
23	7	1.6	306	2 T27985	hypothetical prote
24	7	1.6	310	2 T02280	probable RAV-like
25	7	1.6	313	2 JT0960	polypotein - pota
26	7	1.6	324	2 T23876	hypothetical prote
27	7	1.6	326	2 A43939	proteinase inhibit
28	7	1.6	338	2 F69437	hypothetical prote
29	7	1.6	342	2 G64411	hypothetical prote

30	7	1.6	342	2 C64394	hypothetical prote
31	7	1.6	349	2 T43920	ytic protein [impo
32	7	1.6	357	2 A45619	ubiquinol--cythro
33	7	1.6	370	2 G83219	probable dihydroli
34	7	1.6	379	2 S13556	genome polypotein
35	7	1.6	381	2 G82132	ribonuclease D WC1
36	7	1.6	382	2 S13237	tail fiber protein
37	7	1.6	384	2 S70638	tubulin alpha chai
38	7	1.6	384	2 S70641	tubulin alpha chai
39	7	1.6	384	2 S70640	tubulin alpha chai
40	7	1.6	387	2 S33517	tubulin alpha chai
41	7	1.6	407	2 T19895	hypothetical prote
42	7	1.6	415	2 H70477	UDP-N-acetylmurmo
43	7	1.6	419	2 F85064	hypothetical prote
44	7	1.6	419	2 T25377	hypothetical prote
45	7	1.6	423	1 UBFA	tubulin alpha-1 ch

ALIGNMENTS

RESULT 1
E81195
phospholipase A1, probable NMB0464 [imported] - Neisseria meningitidis (strain MC58 s
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: E81195
R:Telletlin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
ri, H.; Qiu, H.; Yamathayan, J.; Gill, J.; Scariato, V.; Maignani, V.; Piazza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappunli, R.;
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755
A:Accession: E81195
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-382 <TER>
A:Cross-references: GB:AE02403; GB:AE002098; NID:g7225688; PID:NAF40901.1; PID:g722
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0464

Query Match 2.0%, Score 9; DB 2; Length 382;
Best Local Similarity 100.0%; Pred. No. 0.51;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 279 PEIPLTPV.287
Db 221 PEIPLTPV.229

RESULT 2
H81831
probable phospholipase NMA2021 [imported] - Neisseria meningitidis (strain Z2491 serog
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491
A:Reference number: A81775; MUID:20222556
A:Accession: H81831
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-409 <PAR>
A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PID:CA85240.1; PID:g738
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA2021

Query Match 2.0%; Score 9; DB 2; Length 409;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 PEFLTPPV 287
|||||
Db 248 PEFLTPPV 256

RESULT 3
B36971

outer membrane phospholipase A (EC 3.1.1.-) precursor - *Klebsiella pneumoniae*
C:Species: *Klebsiella pneumoniae*
C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 18-Jun-1999
C:Accession: B36971; S40129
R:Brok, R.G.P.M.; Brinkman, E.; van Boxtel, R.; Bekkers, A.C.A.P.A.; Verheij, H.M.; Tomm
J. Bacteriol. 176, 861-870, 1994
A:Title: Molecular characterization of enterobacterial *pldA* genes encoding outer membran
A:Reference number: A36971; MUID:94131966
A:Accession: B36971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-286 <BRO>
A:Cross-references: EMBL:X76901; NID:g436880; PIDN:CAA54223.1; PID:g436881
A:Note: authors translated the codon AAG for residue 112 as Arg
C:Genetics:
A:Gene: *pldA*
C:Superfamily: bacterial phospholipase A1
C:Keywords: carboxylic ester hydrolase

Query Match 1.8%; Score 8; DB 2; Length 286;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 431 GVGIMLND 438
|||||
Db 277 GVGIMLND 284

RESULT 4
T47989
RAY-like protein - *Arabidopsis thaliana*
N:Alternate names: protein F21F14.140
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47989
R:Cholet, N.; Robert, C.; Brottier, P.; Wincker, P.; Catolico, L.; Artiguenave, F.; Se
submitted to the Protein Sequence Database, February 2000
A:Reference number: Z24481
A:Accession: T47989
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-299 <CHO>
A:Cross-references: EMBL:AL138642
A:Experimental source: cultivar Columbia; BAC clone F21F14
C:Genetics:
A:Map position: 3
A:Note: F21F14.140

Query Match 1.8%; Score 8; DB 2; Length 299;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 SLSPDLD 181
|||||
Db 292 SLSPDLD 299

RESULT 5
S36851

L-shaped tail fiber protein - phage T5
N:Alternate names: *ltf* protein
C:Species: phage T5
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 17-Nov-2000
C:Accession: S65934; S01984; S36851

R:Kallman, A.V.; Kulshin, V.E.; Shlyupnikov, M.G.; Ksenzenko, V.N.; Kryukov, V.M.
FEBS Lett. 366, 46-48, 1995
A:Title: The nucleotide sequence of the bacteriophage T5 *ltf* gene.
A:Reference number: S65934; MUID:95309401
A:Accession: S65934

A:Molecule type: DNA
A:Residues: 1-1396 <KAL>
A:Cross-references: EMBL:X69460; NID:g15415; PIDN:CAA9220.1; PID:g15416
R:Kallman, A.V.; Kryukov, V.M.; Bayev, A.A.
Nucleic Acids Res. 16, 6230, 1988
A:Title: The nucleotide sequence of bacteriophage T5 DNA at the region between early
A:Reference number: S01982; MUID:88289370
A:Accession: S01984
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 934-985, 'A', 987-1396 <KA2>
A:Cross-references: EMBL:X07559
C:Genetics:
A:Gene: *ltf*
C:Keywords: late protein; tail fiber

Query Match 1.8%; Score 8; DB 2; Length 1396;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 54 QSATQSAS 61
|||||
Db 87 QSATQSAS 94

RESULT 6
E75377
hypothetical protein - *Deinococcus radiodurans* (strain R1)
C:Species: *Deinococcus radiodurans*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: E75377
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
M.; Shen, M.; Yamathavan, J.J.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
A:Reference number: A75250; MUID:20036896
A:Accession: E75377
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-122 <WHI>
A:Cross-references: GB:AE002002; GB:AE00513; NID:g6459345; PIDN:AAF1150.1; PID:g645
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1580
A:Map position: 1

Query Match 1.6%; Score 7; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 EHEPELY 76
|||||
Db 115 EHEPELY 121

RESULT 7
A29699
beta-lactoglobulin - eastern gray kangaroo
C:Species: *Macropus giganteus* (eastern gray kangaroo)
C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 18-Jul-1997

C:Accession: A29699
R:Godovac-Zimmermann, J.; Shaw, D.
Biol. Chem. Hoppe-Seyler 368, 879-886, 1987
A:Title: The primary structure, binding site and possible function of beta-lactoglobulin
A:Reference number: A29699; MUID:87299024
A:Accession: A29699
A:Molecule type: protein
A:Residues: 1-155 <GOD>
C:Superfamily: lipocalin; lipocalin homology
F:10-155/Domain: lipocalin homology <LIP>

Query Match 1.6%; Score 7; DB 2; Length 155;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 38 RSKNDLG 44
|||||||
DB 5 RSKNDLG 11

RESULT 8
E86498
pts IIA Protein [imported] - Chlamydomonas pneumoniae (strain J138)
C:Species: Chlamydomonas pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: E86498
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349
A:Accession: E86498
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-158 <STO>
A:Cross-references: GB:BA000008; NID:98978434; PIDN:BAA98271.1; GSPDB:GN00142
C:Experimental source: strain J138
C:Genetics:
A:Gene: ptnN_1
C:Superfamily: phosphotransferase system enzyme II; phosphotransferase system mannitol-8

Query Match 1.6%; Score 7; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 LSTLTLS 11
|||||||
DB 126 LSTLTLS 132

RESULT 9
E72124
pts IIA protein - Chlamydomonas pneumoniae (strain CML029)
C:Species: Chlamydomonas pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
C:Accession: E72124
R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606
A:Accession: E72124
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-158 <ARN>
A:Cross-references: GB:AE001591; GB:AE001363; NID:94376311; PIDN:AAD18213.1; PID:9437631
C:Experimental source: strain CML029
C:Genetics:
A:Gene: ptnN_1
C:Superfamily: phosphotransferase system enzyme II; phosphotransferase system mannitol-8
F:28-156/Domain: phosphotransferase system mannitol-specific enzyme II factor III homolog

Query Match 1.6%; Score 7; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 LSTLTLS 11
|||||||
DB 126 LSTLTLS 132

RESULT 10
A54506
tubulin alpha-1 chain - Plasmodium yoelii (fragment)
C:Species: Plasmodium yoelii
C:Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 13-Aug-1999
C:Accession: A54506
R:Arkella, R.; Arasu, P.; Vaidya, A.B.
Mol. Biochem. Parasitol. 30, 165-174, 1988
A:Title: Molecular clones of alpha-tubulin genes of Plasmodium yoelii reveal an unus
A:Reference number: A54506; MUID:89014607
A:Accession: A54506
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-197 <AKE>
A:Cross-references: GB:M29816; NID:g160729; PIDN:AAA29779.1; PID:g160730
C:Superfamily: tubulin

Query Match 1.6%; Score 7; DB 2; Length 197;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120
|||||||
DB 82 IKTKRSI 88

RESULT 11
T19438
hypothetical protein C25A1.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T19438
R:Mortimore, B.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19124
A:Accession: T19438
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-219 <WIL>
A:Cross-references: EMBL:Z81038; PIDN:CAB02762.1; GSPDB:GN00019; CESP:C25A1.1
A:Experimental source: clone C25A1
C:Genetics:
A:Gene: CESP:C25A1.1
A:Map position: 1
A:Introns: 33/1; 66/1; 151/2

Query Match 1.6%; Score 7; DB 2; Length 219;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 158 KDAKOLE 164
|||||||
DB 52 KDAKOLE 58

RESULT 12
A61344
tubulin alpha chain - common octopus (fragment)
C:Species: Octopus vulgaris (common octopus)
C:Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 10-Jul-1998
C:Accession: A61544
R:Zinov'eva, R.D.; Aleinikova, K.S.; Tomarev, S.I.

Dokl. Akad. Nauk SSSR 302, 462-467, 1988

A:Title: Isolation and structural characterization of cDNAs coding for alpha-tubulin of A:Reference number: A61544

A:Accession: A61544

A>Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-240 <2IN>

C:Superfamily: tubulin

F:239-240/Cleavage site: Glu-Tyr (tubulin-specific carboxypeptidase) #status predicted

F:239-240/Cross-Link: peptide (Glu-Tyr) (by tubulin-tyrosine ligase) #status predicted

Query Match 1.6%; Score 7; DB 2; Length 240;

Best Local Similarity 100.0%; Pred. No. 42;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTAKSI 120

Db 124 IKTAKSI 130

RESULT 13

A44959

coat protein - potato virus Y (strain N) (fragment)

C:Species: potato virus Y, PVY

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Nov-2000

C:Accession: A44959

R:Hay, J.M.; Fellowes, A.P.; Timmerman, G.M.

Arch. Virol. 107, 111-122, 1989

A:Title: Nucleotide sequence of the coat protein gene of a necrotic strain of potato virus Y

A:Reference number: A44959; MID:90025730

A:Accession: A44959

A>Status: Preliminary

A:Molecule type: genomic RNA

A:Residues: 1-264 <HAY>

A:Cross-references: GB:M2470; NID:9333508; PIDN:AAA47182.1; PID:9333509

C:Superfamily: tobacco etch virus genome polypeptide

Query Match 1.6%; Score 7; DB 2; Length 264;

Best Local Similarity 100.0%; Pred. No. 45;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 TKKDAKQ 162

Db 8 TKKDAKQ 14

RESULT 14

A60366

coat protein - potato virus Y (strain TH)

C:Species: potato virus Y, PVY

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Nov-2000

C:Accession: A60366

R:Hataya, T.; Sano, T.; Ohshima, K.; Shikata, E.

Virus Genes 4, 339-350, 1990

A:Title: Polymerase chain-reaction-mediated cloning and expression of the coat protein gene of potato virus Y

A:Reference number: A60366; MID:91143125

A:Accession: A60366

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-267 <HAT>

C:Superfamily: tobacco etch virus genome polypeptide

Query Match 1.6%; Score 7; DB 2; Length 267;

Best Local Similarity 100.0%; Pred. No. 46;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 TKKDAKQ 162

Db 11 TKKDAKQ 17

RESULT 15

S14001

genome polypeptide - potato virus Y (strain H) (fragment)

N:Contains: coat protein

C:Species: potato virus Y, PVY

A:Variety: strain H

C:Date: 21-Nov-1993 #sequence_revision 08-Nov-1996 #text_change 17-Nov-2000

C:Accession: S14001

R:Daimy, T.; Balazs, E.

Nucleic Acids Res. 18, 6721, 1990

A:Title: Nucleotide sequence of an altered virulence potato virus Y coat protein gene

A:Reference number: S14001; MID:91067494

A:Accession: S14001

A>Status: translation not shown

A:Molecule type: genomic RNA

A:Residues: 1-267 <DAL>

A:Cross-references: EMBL:X54611; NID:961449; PIDN:CAA38432.1; PID:9330257

A:Experimental source: strain H

C:Superfamily: tobacco etch virus genome polypeptide

F:1-267/Product: coat protein #status predicted <CPR>

Query Match 1.6%; Score 7; DB 2; Length 267;

Best Local Similarity 100.0%; Pred. No. 46;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 TKKDAKQ 162

Db 11 TKKDAKQ 17

Search completed: May 9, 2002, 12:56:33

Job time: 603 sec

1	8	1.8	286	PAL_KLEBN
2	8	1.8	1396	VLTB_BPp5
3	8	1.6	155	LACB_MAGCI
4	7	1.6	174	LACB_MACCEU
5	7	1.6	186	PRL_ICPpU
6	7	1.6	197	TBA_PIAVO
7	7	1.6	240	TBA_OCTYU
8	7	1.6	270	Y304_MYCPN
9	7	1.6	273	ILVE_THEMA
10	7	1.6	322	MIAA_PSEPU
11	7	1.6	326	SPI_BACBR
12	7	1.6	342	FLAD_METJA
13	7	1.6	342	Y755_METJA
14	7	1.6	423	AMT_AOUJE
15	7	1.6	444	TBA_ONCKE
16	7	1.6	448	TBA4_HUBAN
17	7	1.6	448	TBA5_CHICK
18	7	1.6	449	TBA2_DROME
19	7	1.6	449	TBAE_PHYPO
20-	7	1.6	449	TBAN_PHYPO
21	7	1.6	449	TBA_XENLA
22	7	1.6	450	TBAD_PHYPO
23	7	1.6	450	TBA_HABCO
24	7	1.6	450	TBA_NOTYT
25	7	1.6	451	TBA1_CRIGR
26	7	1.6	451	TBA3_HOMAM
27	7	1.6	451	TBA_OCTDO
28	7	1.6	451	TBA_TORNA
29	7	1.6	511	LGT_CITUN
30	7	1.6	513	ATPA_HARIN
31	7	1.6	536	GOIM_BUCAF
32	7	1.6	544	GPI0_DICTDI
33	7	1.6	567	GVP_RAT

34	7	1.6	610	1	F1MB_DIC01
35	7	1.6	635	1	DNAR_HA1MA
36	7	1.6	768	1	YR23_HUMN
37	7	1.6	783	1	FYE_HUMAN
38	7	1.6	900	1	AXN1_HUMAN
39	7	1.6	947	1	LKTA_PASSP
40	7	1.6	1026	1	VG37_BP7A
41	7	1.6	3061	1	POLG_PVYHU
42	6	1.4	33	1	BR2E_RANES
43	6	1.4	33	1	BR2E_RANES
44	6	1.4	50	1	YR30_CAEBL
45	6	1.4	51	1	INS_ACOCA
					P01324
					acnoms canh
					caenornabdt
					halocostelli
					P05460
					P01100
					halocostelli
					O9u117
					homo sapien
					O15117
					O15169
					homo sapien
					P05123
					pasteurella
					P03744
					bacteriophla
					Q002963
					p genome po
					P040838
					rana esculat
					P432413
					rana esculat
					P33001
					caenornabdt
					P01324
					acnoms canh

ALIGNMENTS

	RESULT	1			
ID	PAL_KLEPN	STANDARD;	PRT:	286 AA.	
AC	p37446:				
DT	01-OCT-1994 (Rel. 30, Created)				
DT	01-OCT-1994 (Rel. 30, Last sequence update)				
DT	01-NOV-1997 (Rel. 35, Last annotation update)				
DE	PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETURGENT-RESISTANT				
DE	PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPATIDYLCHOLINE 1-				
DE	ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLLIPASE A) (OM PLA).				
CN	PLA..				
OS	Klebsiella pneumoniae.				
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;				
CC	Klebsiella.				
OX	NCBI_TaxId=573;				
RN	[1]				
RP	SEQUENCE FROM N.A..				
RX	MEDLINE=94131966; PubMed=8300539;				
RA	Brok R.G.P.M., Brinkman E., van Boxtel R., Bekkers A.C.A.P.,				
RA	Verheij H.M., Tommassen J.;				
RT	"Molecular characterization of enterobacterial pla genes encoding				
RT	outer membrane phospholipase A.";				
RL	J. Bacteriol. 176:861-870(1994).				
CC	-1- FUNCTION: HYDROLYSIS OF PHOSPATIDYLCHOLINE WITH PHOSPHOLIPASE				
CC	A2 (EC 3.1.1.14) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.				
CC	-1- CATALYTIC ACTIVITY: PHOSPATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-				
CC	-1- CATALYTIC ACTIVITY: A FATTY ACID ANION.				
CC	-1- CATALYTIC ACTIVITY: PHOSPATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-				
CC	-1- CATALYTIC ACTIVITY: PHOSPATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-				
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.				
CC	-1- COPFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.				
CC	-1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES				
CC	LOCATED THERE.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way com-				
CC	mercialized and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
CC	EMBL; X76901; CAAS4223.1; ..				
DR	PIR; B36971; B36971.				
DR	PIR; S40129; S40129.				
DR	InterPro: IPRO03187; PLAI.				
DR	Pfam: PF02253; PLAI; 1.				
KW	Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium.				
FT	SIGNAL	1	20		
FT	CHAIN	21	286		
FT	ACT_SITE	161	161		
SQ	SEQUENCE	286 AA;	32544 MW;	3E39F86308510BA3 CRC64;	
	Query Match	1.8%;	Score 8;	DB 1;	Length 286;
	Best Local Similarity	100.0%;	Pred. No. 1.6;		
	Matches	8;	Conservative 0;	Mismatches 0;	Gaps 0;

OY 431 GVGIMLND 438
 DB 277 GVGIMLND 284

RESULT 2

VLTF_BPRT5 STANDARD: PRT; 1396 AA.
 AC P13390; 048502;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE L-SHAPED TAIL FIBER PROTEIN (LTF PROTEIN).
 GN LTF.
 OS Bacteriophage T5.
 OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae.
 OX NCBI_TaxID=10726;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95309401; PubMed=7789514;
 RA Kallman A.V., Kulshin V.E., Shlyapnikov M.G., Keenzenko V.N.,
 RA Kryukov V.M.;
 RT "The nucleotide sequence of the bacteriophage T5 ltf gene.";
 RL FEBS Lett. 366:46-48(1995).
 RN (2)
 RP SEQUENCE FROM N.A.
 RA Kallman A.V.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN (3)
 RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
 RX MEDLINE=88289370; PubMed=3267228;
 RA Kallman A.V., Kryukov V.M., Bayev A.A.;
 RT "The nucleotide sequence of bacteriophage T5 DNA at the region
 RT between early and late genes.";
 RL Nucleic Acids Res. 16:6230-6230(1988).
 CC -1- FUNCTION: NONESSENTIAL PROTEIN THAT MEDIATES BINDING TO THE
 CC POLYMANNOSE O ANTIGEN.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X69460; CAA49220.1; -.
 DR EMBL: AJ001191; CAA04591.1; -.
 DR PIR: S01982; S01982.
 KW Late protein.
 FT CONFLICT 986 V -> A (IN REF. 2).
 SQ SEQUENCE 1396 AA; 147989 MW; 18CD2192F65FFC1 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 1396;
 Best Local Similarity 100.0%; Pred. No. 6.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 54 QSATOSAS 61
 DB 87 QSATOSAS 94

RESULT 3

LACB_MACGI STANDARD: PRT; 155 AA.
 AC P11944;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE BETA-LACTOGLOBULIN.
 GN LGB.

OS Macropus giganteus (Eastern gray kangaroo).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
 OX NCBI_TaxID=9317;
 RN (1)
 RP SEQUENCE.

RX MEDLINE=87299024; PubMed=3620116;
 RA Godovac-Zimmermann J., Shaw D.;
 RT "Beta-lactoglobulin identified in marsupial milk. The primary
 RT structure, binding site and possible function of beta-lactoglobulin
 RT from eastern grey kangaroo (Macropus giganteus).";
 RL Biol. Chem. Hoppe-Seyler 368:879-886(1987).
 CC -1- FUNCTION: LACTOGLOBULIN IS THE PRIMARY COMPONENT OF WHEY, IT
 CC BINDS RETINOL AND IS PROBABLY INVOLVED IN THE TRANSPORT OF
 CC THAT MOLECULE.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SIMILARITY: BELONGS TO THE LIPOCALIN FAMILY.
 DR PIR: A29699; A29699.
 DR HSSP: P02754; 1BSQ.
 DR InterPro: IPR000566; Lipocln_cytfabp.
 DR Pfam: PF00061; Lipocalin.1.
 DR PROSITE: PS00213; LIPOCALIN: FALSE NEG.
 KW Milk; whey; Retinol-binding; Transport; Lipocalin.
 FT DISULFID 61 153 BY SIMILARITY.
 FT DISULFID 103 115
 SQ SEQUENCE 155 AA; 18340 MW; 7BD7EA4A191530E5 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 155;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 38 RSKNDLG 44
 DB 5 RSKNDLG 11

RESULT 4

LACB_MACEU STANDARD: PRT; 174 AA.
 AC Q29614;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE BETA-LACTOGLOBULIN PRECURSOR.
 GN LGB OR BLG.
 OS Macropus eugenii (Tamar wallaby).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
 OX NCBI_TaxID=9315;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91197370; PubMed=1707629;
 RA Collet C., Joseph R., Nicholas K.R.;
 RT "A marsupial beta-lactoglobulin gene: characterization and prolactin-
 RT dependent expression.";
 RL J. Mol. Endocrinol. 6:9-16(1991).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95314595; PubMed=7794241;
 RA Collet C., Joseph R.;
 RT "Exon organization and sequence of the genes encoding alpha-
 RT lactalbumin and beta-lactoglobulin from the tamar wallaby
 RT (Macropodidae, Marsupialia).";
 RL Biochem. Genet. 33:61-72(1995).
 RN (3)
 RP SEQUENCE OF 47-174 FROM N.A.
 RC TISSUE=Mammary gland;
 RX MEDLINE=95085593; PubMed=7993373;
 RA Collet C., Joseph R.;
 RT "The identification of nuclear and mitochondrial genes by sequencing
 RT randomly chosen clones from a marsupial mammary gland cDNA library.";
 RL Biochem. Genet. 32:181-190(1994).

CC -1- FUNCTION: LACTOGLOBULIN IS THE PRIMARY COMPONENT OF WHEY, IT
 CC BINDS RETINOL AND IS PROBABLY INVOLVED IN THE TRANSPORT OF
 CC THAT MOLECULE.
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE LIPOCALIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L14959; AAA31597.1; -.
 DR EMBL: L14954; AAA31597.1; JOINED.
 DR EMBL: L14955; AAA31597.1; JOINED.
 DR EMBL: L14956; AAA31597.1; JOINED.
 DR EMBL: L14957; AAA31597.1; JOINED.
 DR EMBL: L14958; AAA31597.1; JOINED.
 DR EMBL: X15212; CAA33282.1; -.
 DR HSSP: P02754; IBSO.
 DR InterPro: IPR000566; Lipocln_cyFABP.
 DR Pfam: PF00061; Lipocalin; 1.
 DR PROSITE: PS00213; LIPOCALIN; FALSE NEG.
 DR Milk: Whey; Retinol-binding; Transport; Lipocalin; Signal.
 KW SIGNAL
 FT CHAIN 1 18
 FT DISULEID 19 174
 FT DISULEID 79 172
 FT DISULEID 122 134
 FT CONFLICT 76 76 A -> T (IN REF. 3).
 SQ SEQUENCE 174 AA; 20219 MW; 6C5901051F99C991 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 174;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 RSKNDLG 44
 Db 23 RSKNDLG 29
 PRL_ICTPU 5
 ID PRL_ICTPU STANDARD; PRT; 186 AA.
 AC P51904;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE PROLACTIN (PRL).
 OS Ictalurus punctatus (Channel catfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Siluriformes; Ictaluridae; Ictalurus.
 OX NCBI_TaxID=7998;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Pituitary;
 RX MEDLINE=93364578; PubMed=1308206;
 RA Watanabe K., Igarashi A., Noso T., Chen T.T., Dunham R.A.,
 RA Kawachi H.;
 RT "Chemical identification of catfish growth hormone and prolactin."
 RL Mol. Mar. Biol. Biotechnol. 1:239-249(1992).
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
 DR HSSP: Q28633; IAN3.
 DR InterPro: IPR001400; SOMATOTROPIN.
 DR Pfam: PF00103; hormone; 1.
 DR PRINTS: PR00836; SOMATOTROPIN.
 DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
 DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
 KW Hormone; Pituitary.

FT DISULEID 45 159 BY SIMILARITY.
 FT DISULEID 176 186 BY SIMILARITY.
 SQ SEQUENCE 186 AA; 20606 MW; 7BCCD89718F44E74 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 186;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 NDLGQDN 47
 Db 140 NDLGQDN 146

RESULT 6
 TBA_PLAYO
 ID TBA_PLAYO STANDARD; PRT; 197 AA.
 AC P12543;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE TUBULIN ALPHA CHAIN (FRAGMENT).
 OS Plasmodium berghei yoelii.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5862;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89014607; PubMed=2459618;
 RA Akella R., Arasu P., Valdivia A.B.;
 RT "Molecular clones of alpha-tubulin genes of Plasmodium yoelii reveal
 RT an unusual feature of the carboxy terminus."
 RL Mol. Biochem. Parasitol. 30:165-174(1988).
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
 CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
 CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
 CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M29816; AAA29779.1; -.
 DR PIR: A54506; A54506.
 DR InterPro: IPR000217; Tubulin.
 DR InterPro: IPR003008; Tubulin_FtsZ.
 DR Pfam: PF00091; tubulin; 1.
 DR PROSITE: PS00227; TUBULIN; PARTIAL.
 KW Microtubules; GTP-binding.
 FT NON TER 1
 SQ SEQUENCE 197 AA; 22111 MW; 4DAAF199CCGCD319 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 197;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 IKTKRSI 120
 Db 82 IKTKRSI 88

RESULT 7
 TBA_OCTVU
 ID TBA_OCTVU STANDARD; PRT; 240 AA.
 AC P24635;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE TUBULIN ALPHA CHAIN (FRAGMENT).
 OS Octopus vulgaris (Octopus).
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;
 OC Incirrata; Octopodidae; Octopus.
 OX NCBI_TaxID=6645;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lens;
 RA Zinov'eva R.D., Aleinikova K.S., Tomarev S.I.;
 RT "Isolation and structural characterization of cDNAs coding for alpha-
 RT tubulin of the octopus eye lens."
 RL Dokl. Akad. Nauk SSSR 302:462-467(1988).
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
 CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
 CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
 CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to [license@isb-sid.ch](mailto:license@isb-sib.ch)).
 CC -----
 CC EMBL: X15845; CA33844.1; -
 CC DR PIR: A61544; A61544.
 CC DR InterPro: IPR000217; Tubulin.
 CC DR InterPro: IPR003008; Tubulin_FtsZ.
 CC DR Pfam: PF00091; tubulin; 1.
 CC DR PROSITE: PS00227; TUBULIN; PARTIAL.
 CC KW Microtubules; GTP-binding.
 CC FT NON_TER 1
 CC FT SITE 240 240 INVOLVED IN POLYMERIZATION.
 CC SO SEQUENCE 240 AA; 26961 MW; 13BB3A1F740F2A16 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 240;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 114 IKTKRSI 120
 Db 124 IKTKRSI 130

RESULT 8
 ID Y304_MYCPN STANDARD; PRT; 270 AA.
 AC P75355;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN MG304 HOMOLOG
 DE (A05_ORF270L).
 GN MPN433 OR MP408.
 OS Mycoplasma pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molluscites;
 OC Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RX MEDLINE=97105885; Pubmed=8948633;
 RA Himmelfeich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
 RA Herrmann R.;
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
 RT pneumoniae."
 RT Nucleic Acids Res. 24:4420-4449(1996).
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 CC (ABC TRANSPORTERS).
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to [license@isb-sid.ch](mailto:license@isb-sib.ch)).
 CC -----
 CC EMBL: AE000040; AAB96056.1; -
 CC DR InterPro: IPR003593; AAA.
 CC DR InterPro: IPR003439; ABC_transporter.
 CC DR InterPro: IPR001687; ATP_GTP_A.
 CC DR Pfam: PF00005; ABC_tran; 1.
 CC DR SMART: SM00382; AAA; 1.
 CC DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 CC KW Hypothetical protein; ATP-binding; Transport; Complete proteome.
 CC FT NP_BIND 36 43
 CC FT SEQUENCE 270 AA; 30770 MW; 6512640E4BC051B4 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 270;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 168 KQFTPLS 174
 Db 125 KQFTPLS 131

RESULT 9
 ID LIVE_THEMA STANDARD; PRT; 273 AA.
 AC P74921;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PROBABLE BRANCHED-CHAIN AMINO ACID AMINOTRANSFERASE (EC 2.6.1.42)
 DE (BCAT).
 GN LIVE OR TM0831.
 OS Thermotoga maritima.
 OC Bacteria; Thermotogales; Thermotoga.
 OX NCBI_TaxID=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=99287316; Pubmed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
 RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Utterback T.R., Malek J.A., Linner K.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RT genome sequence of Thermotoga maritima."
 RT Nature 399:323-329(1999).
 RN [2]
 RP SEQUENCE OF 68-273 FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=97017137; Pubmed=8863738;
 RA Gulpaud O., Labedan B., Forterre P.;
 RT "A gyrB-like gene from the hyperthermophilic bacterium Thermotoga
 RT maritima."
 RT Gene 174:121-128(1996).
 CC -1- CATALYTIC ACTIVITY: L-LEUCINE + 2-OXOGLUTARATE = 4-METHYL-2-
 CC OXOPENTANOATE + L-GLUTAMATE (ALSO ACTS ON L-ISOLEUCINE AND
 CC L-VALINE).
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -1- PATHWAY: VALINE AND ISOLEUCINE BIOSYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO CLASS-IV OF PYRIDOXAL-PHOSPHATE-DEPENDENT
 CC AMINOTRANSFERASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AE001750; AAD35913.1; -
 DR EMBL: U96692; AAC4497.1; -
 DR HSSP: P00510; 1A3G.
 DR TIGR: TM0831; -
 DR InterPro: IPR001544; Aminotran_4.
 DR Pfam: PF01063; aminotran_4; 1.
 DR ProDom: PD001961; Aminotran_4; 1.
 DR PROSITE: PS00770; AA_TRANSFERR_CLASS_4; 1.
 DR Transferrase; Aminotransferase; Branched-chain amino acid biosynthesis;
 KW Pyridoxal phosphate; Complete proteome.
 FT CONFLICT 77 A -> R (IN REF. 2).
 FT SEQUENCE 273 AA; 31158 MW; 21628705612E90E3 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 273;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 172 PLSLSPD 178
 Db 61 PLSLSPD 67

RESULT 10
 MIAA_PSEPU STANDARD; PRT; 322 AA.
 AC 030762;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE TRNA DELTA(2) -ISOPENTENYL-PYROPHOSPHATE TRANSFERASE (EC 2.5.1.8) (TPP
 DE TRANSFERASE) (ISOPENTENYL-DIPHOSPHATE:TRNA ISOPENTENYLTRANSFERASE)
 DE (TPPASE) (TPP).
 GN MIAA.
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M.
 RA Olekhnovich I.N., Gussin G.N.;
 RT "Attenuation of the Pseudomonas putida type and trpGDC genes.";
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CATALYZES THE FIRST STEP IN THE BIOSYNTHESIS OF
 CC 2-METHYLTHIO-N6-(DELTA(2)-ISOPENTENYL)-ADENOSINE (MS[2]116[a])
 CC ADJACENT TO THE ANTICODON OF SEVERAL TRNA SPECIES (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ISOPENTENYL DIPHOSPHATE + TRNA =
 CC PYROPHOSPHATE + TRNA CONTAINING 6-ISOPENTENYLADENOSINE.
 CC -1- SIMILARITY: BELONGS TO THE TPP TRANSFERASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL: AF016312; AAB69443.1; -
 DR InterPro: IPR002627; TPPt.
 DR Pfam: PF01715; TPPt; 1.
 DR ProDom: PD004674; TPPt; 1.
 DR Transferrase; Nucleotidyltransferase; RNA processing; ATP-binding.
 KW NE_BIND 12 19 ATP (POTENTIAL).
 FT SEQUENCE 322 AA; 35488 MW; D1AA81A2D7B32F6A CRC64;

Query Match 1.6%; Score 7; DB 1; Length 322;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 FVDEVR 39
 Db 231 FVDEVR 237

RESULT 11
 SPL_BACBR STANDARD; PRT; 326 AA.
 AC P43131;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE PROTEASE INHIBITOR PRECURSOR (BBRP).
 DE Bacillus brevis.
 OS Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Brevibacillus.
 OX NCBI_TaxID=1393;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-32; 104-112 AND 122-136.
 RC STRAIN=HPD31;
 RX MEDLINE=92304060; PubMed=1610177;
 RA Shiga Y., Hasegawa K., Tsuboi A., Yamagata H., Ueda S.;
 RT "Characterization of an extracellular protease inhibitor of Bacillus
 RT brevis HPD31 and nucleotide sequence of the corresponding gene.";
 RL Appl. Environ. Microbiol. 58:525-531(1992).
 CC -1- FUNCTION: SHOWS INHIBITORY ACTIVITY TOWARDS SERINE PROTEASES, SUCH
 CC AS TRYPSIN, CHYMOTRYPSIN, AND SUBTILISIN. MAY FORM A TRYPSIN-
 CC INHIBITOR COMPLEX IN A MOLAR RATIO OF 1:1. IT IS HEAT RESISTANT AT
 CC NEUTRAL AND ACIDIC PH.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- PTM: MAY UNDERGO ACTIVATION AFTER SECRETION. IT IS PRODUCED
 CC EXTRACELLULARLY IN MULTIPLE FORMS HAVING AT LEAST THREE DIFFERENT
 CC MOLECULAR WEIGHTS (BBRP-A, -B, AND -C).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL: D10696; BAA01538.1; -
 DR Serine protease inhibitor; Protease inhibitor; Signal; Repeat.
 KW STGNAL 1 24
 FT STGNAL 1 24
 FT CHAIN 25 326 SERINE PROTEASE INHIBITOR.
 FT CHAIN 104 326 SERINE PROTEASE INHIBITOR-C.
 FT CHAIN 104 7 SERINE PROTEASE INHIBITOR-B (POTENTIAL).
 FT CHAIN 122 326 SERINE PROTEASE INHIBITOR-A.
 FT DOMAIN 177 304 CONTAINS TWO APPROXIMATE REPEATS.
 FT REPEAT 177 208 1.
 FT REPEAT 272 304 2.
 SO SEQUENCE 326 AA; 35100 MW; 1C0456BAF912F77 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 326;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 PVAFVDE 36
 Db 119 PVAFVDE 125

RESULT 12
 FLAD_METUA STANDARD; PRT; 342 AA.
 ID FLAD_METUA
 AC Q58305;

```

DT 15-JUL-1998 (rel. 36, Created)
DT 15-DEC-1998 (rel. 37, Last sequence update)
DT 20-AUG-2001 (rel. 40, Last annotation update)
DE PUTATIVE FLAGELLA-RELATED PROTEIN D.
DE FLAD OR MJ0895.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhman J.L., Nguyen D.,
RA Uitterback L.R., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Cotton M.D., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RT Science 273:1058-1073(1996).
CC -1- SIMILARITY: STRONG, TO M.VOLUTAE FLAD, ALSO TO FLAE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
CC EMBL: U67521; AAB89748.1; -.
CC TIGR: MJ0755; -.
CC FLAGELLA: Complete proteome.
CC SEQUENCE 342 AA; 39950 MW; B384DDE1775566C CRC64;
SQ

```

Query Match 1.6%; Score 7; DB 1; Length 342;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 356 PDLIDY 362
DB 255 PDLIDY 261

```

RESULT 13
Y755_METJA STANDARD; PRT; 342 AA.

```

AC Q58165;
DT 01-NOV-1997 (rel. 35, Created)
DT 01-NOV-1997 (rel. 35, Last sequence update)
DT 20-AUG-2001 (rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN MJ0755 PRECURSOR.
GN MJ0755.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhman J.L., Nguyen D.,
RA Uitterback L.R., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,

```

```

RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RT Science 273:1058-1073(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
CC EMBL: U67521; AAB89748.1; -.
CC TIGR: MJ0755; -.
CC HYPOTHETICAL protein; Signal; Complete proteome.
CC SIGNAL 1 18
CC CHAIN 19 342
CC SEQUENCE 342 AA; 38778 MW; BD35220A7EAD85C CRC64;
SQ

```

Query Match 1.6%; Score 7; DB 1; Length 342;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 76 YTTALEN 82
DB 90 YTTALEN 96

```

RESULT 14
AMT_AQUAE STANDARD; PRT; 423 AA.

```

AC Q66515;
DT 30-MAY-2000 (rel. 39, Created)
DT 30-MAY-2000 (rel. 39, Last sequence update)
DT 20-AUG-2001 (rel. 40, Last annotation update)
DE PROBABLE AMMONIUM TRANSPORTER.
GN AMT OR AMTB OR AO_112.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Anjay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus."
RT Nature 392:353-358(1998).
RL -1- FUNCTION: INVOLVED IN THE UPTAKE OF AMMONIA (PROBABLE).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE AMT1/MEP/MRGA FAMILY OF AMMONIUM
CC TRANSPORTERS (TC 2.49).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
CC EMBL: AE00674; AAC06478.1; -.
CC InterPro: IPR001905; Ammonium_transp.
DR Pfam: PF00909; Ammonium_transp. 1.
DR PROSITE: PS01219; AMMONIUM_TRANS 1.
KW transport; Transmembrane; Complete proteome.
FT TRANSMEM 3 23
FT TRANSMEM 31 51
FT TRANSMEM 64 84
FT TRANSMEM 64 84
FT POTENTIAL.
FT POTENTIAL.

```

```

FT  TRANSMEM  123  143  POTENTIAL.
FT  TRANSMEM  150  170  POTENTIAL.
FT  TRANSMEM  183  203  POTENTIAL.
FT  TRANSMEM  217  237  POTENTIAL.
FT  TRANSMEM  246  266  POTENTIAL.
FT  TRANSMEM  278  298  POTENTIAL.
FT  TRANSMEM  301  321  POTENTIAL.
FT  TRANSMEM  331  351  POTENTIAL.
FT  TRANSMEM  374  394  POTENTIAL.
SQ  SEQUENCE  423  AA; 44322 MM; 166203BBB2A8A15B CRC64;

```

```

QY  114  IKTKRSI 120
Db  329  IKTKRSI 335

```

Search completed: May 9, 2002, 13:05:52
Job time: 607 sec

Query Match 1.6%; Score 7; DB 1; Length 423;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY  402  LKKGISG 408
    |||||
Db  104  LKKGISG 110

```

```

RESULT 15
TBA_ONCKE
ID  TBA_ONCKE  STANDARD;  PRT;  444  AA.
AC  P30436;
DT  01-APR-1993 (Rel. 25, Created)
DT  01-APR-1993 (Rel. 25, Last sequence update)
DT  15-JUL-1999 (Rel. 38, Last annotation update)
DE  TUBULIN ALPHA CHAIN.
OS  Oncorhynchus keta (Chum salmon).
OC  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC  Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC  Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX  NCBI_Taxid=8018;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Brain;
RX  MEDLINE=93208376; PubMed=1296820;
RA  Coe I.R., Munro R., Sherwood N.M.;
RT  "Isolation of different brain-specific isoforms of alpha-tubulins
RT  from chum salmon (Oncorhynchus keta).";
RL  DNA Seq. 3:257-262(1992).
CC  -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROFUBULES. IT
CC  BINDS TWO MOLES OF GTP. ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC  CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC  -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
CC  -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X66973; CAA47384.1; -
DR  PIR; S25004; S25004.
DR  InterPro; IPR000217; Tubulin.
DR  InterPro; IPR003008; Tubulin_Fts2.
DR  Pfam; PF00091; tubulin; 1.
DR  PRINTS; PR01161; TUBULIN.
DR  PRINTS; PR01162; ALPHATUBULIN.
DR  PROSITE; PS00227; TUBULIN; 1.
KW  Microtubules; GTP-binding; Multigene family.
FT  NP_BIND 136 142 GTP (POTENTIAL).
FT  SITE 444 444 INVOLVED IN POLYMERIZATION.
FT  SEQUENCE 444 AA; 49314 MM; C2CD75432DD6BAF CRC64;
SQ

```

Query Match 1.6%; Score 7; DB 1; Length 444;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 13:05:17 ; Search time 77.39 Seconds
(Without alignments)
835,410 Million cell updates/sec

Title: US-09-787-083-4
Perfect score: 442
Sequence: 1 MKVSLSTLTLSTLPCFALLA.....YVHEATSPGVGLMDMGL 442

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 473505 seqs, 14627329 residues

Word size : 0

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SPREMBL_17:*

- 1: sp.archaea:*
- 2: sp.bacteria:*
- 3: sp.fungi:*
- 4: sp.human:*
- 5: sp.invertebrate:*
- 6: sp.mammal:*
- 7: sp.mhc:*
- 8: sp.organelle:*
- 9: sp.phage:*
- 10: sp.plant:*
- 11: sp.podent:*
- 12: sp.virus:*
- 13: sp.vertebrate:*
- 14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	2.0	382	2 09K007	09K007 neisseria m
2	9	2.0	409	2 09J721	09J721 neisseria m
3	9	2.0	862	2 09CL62	09CL62 pasteurella
4	9	2.0	1432	10 09AX85	09AX85 oryza sativ
5	8	1.8	299	10 09M268	09M268 arabidopsis
6	8	1.8	306	2 09CL22	09CL22 pasteurella
7	7	1.6	55	12 067608	067608 tomato gold
8	7	1.6	72	12 091RS5	091RS5 potato viru
9	7	1.6	72	12 091RS4	091RS4 potato viru
10	7	1.6	95	12 091RS3	091RS3 potato viru
11	7	1.6	96	12 091RS8	091RS8 potato viru
12	7	1.6	101	12 091RS6	091RS6 potato viru
13	7	1.6	122	2 09RU16	09RU16 delinococcus
14	7	1.6	124	2 091RS7	091RS7 potato viru
15	7	1.6	131	6 09GMD4	09GMD4 macaca fusc
16	7	1.6	138	12 091XQ7	091XQ7 potato viru
17	7	1.6	147	5 036035	036035 hexamita in
18	7	1.6	158	2 0929C1	0929C1 chlamydia p
19	7	1.6	169	2 0912A7	0912A7 streptomyces

20	7	1.6	186	10 09LX9	09LX9 arabidopsis
21	7	1.6	189	13 09DFT2	09DFT2 notochenia
22	7	1.6	198	4 09BUX9	09BUX9 homo sapien
23	7	1.6	212	13 091819	091819 ictalurus p
24	7	1.6	216	2 09X6H6	09X6H6 streptococc
25	7	1.6	218	2 09RF14	09RF14 streptococc
26	7	1.6	219	5 017594	017594 caenorhabdi
27	7	1.6	220	2 09XCL1	09XCL1 streptococc
28	7	1.6	220	5 036036	036036 spirocheteu
29	7	1.6	221	12 065002	065002 allia mos
30	7	1.6	227	5 09BMP4	09BMP4 glossina mo
31	7	1.6	233	6 09MZB0	09MZB0 ovis aries
32	7	1.6	236	13 09DPS8	09DPS8 notochenia
33	7	1.6	264	12 085260	085260 potato viru
34	7	1.6	265	9 038155	038155 bacterioph
35	7	1.6	267	9 099363	099363 bacterioph
36	7	1.6	267	12 085261	085261 potato viru
37	7	1.6	267	12 098628	098628 potato viru
38	7	1.6	267	12 086842	086842 potato viru
39	7	1.6	267	12 099026	099026 potato viru
40	7	1.6	267	12 005476	005476 potato viru
41	7	1.6	267	12 009711	009711 potato viru
42	7	1.6	267	12 09WH17	09WH17 potato viru
43	7	1.6	267	12 09QNM9	09QNM9 potato viru
44	7	1.6	267	12 09DUL0	09DUL0 potato viru
45	7	1.6	267	12 09DUX9	09DUX9 potato viru

ALIGNMENTS

RESULT	1	PRELIMINARY	PRT	382 AA
09K007	09K007			
ID	09K007			
AC	09K007			
DT	01-OCT-2000 (TREMBLREL.15, Created)			
DT	01-OCT-2000 (TREMBLREL.15, Last sequence update)			
DT	01-JUN-2001 (TREMBLREL.17, Last annotation update)			
DE	PHOSPHOLIPASE A1, PUTATIVE.			
GN	NMB0464.			
OS	Neisseria meningitidis (serogroup B).			
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.			
OX	NCBI_TaxID=491;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MC58 / SEROGROUP B;			
RX	MEDLINE=20175755; PubMed=10710307;			
RA	Tetelijn H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,			
RA	Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,			
RA	Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey E.K.,			
RA	Halt D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,			
RA	Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,			
RA	Cotton M.D., Uterback T.R., Khouli H., Qin H., Yamathavan J.,			
RA	Gill J., Scarlato V., Maignani V., Piza M., Grandi G., Sun L.,			
RA	Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;			
RT	"Complete genome sequence of Neisseria meningitidis serogroup B strain			
RT	MC58."			
RL	Science 287:1809-1815(2000).			
DR	EMBL: AE002403; AAF40901.1;			
DR	TIGR: NMB0464;			
DR	InterPro: IPR003187; Pfam:			
DR	Pfam: PF02253; Pfam: 1.			
KW	Complete proteome.			
SO	SEQUENCE 382 AA; 42714 MW; B46BA802F062E836 CRC64;			

Query Match	2.0%	Score 9:	DB 2:	Length 382;
Best Local Similarity	100.0%	Pred. No. 1.4;		
Matches 9;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 279	PERFUTOPV 287			
DB 221	PERFUTOPV 229			

```

RESULT 2
Q9JUT21
ID 09JUT21 PRELIMINARY: PRT: 409 AA.
AC 09JUT21:
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE PUTATIVE PHOPHOLIPASE.
GN NMA2021.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBL_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holtroyd S.,
RA Jagers K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491."
RL Nature 404:502-506(2000).
DR EMBL: AL162757; CAB85240.1;
DR InterPro: IPR003187; Pfam:
DR Pfam: PF02253; Pfam: 1.
KW Complete proteome.
SQ SEQUENCE 409 AA; 45862 MW; CD6585B064D01A41 CRC64;

Query Match
Best Local Similarity 2.0%; Score 9; DB 2; Length 409;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIFLTQPV 287
DB 248 PEIFLTQPV 256

RESULT 3
Q9CL62
ID 09CL62 PRELIMINARY: PRT: 862 AA.
AC 09CL62:
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE CVGS1.
GN CVGS1 OR PM1380.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBL_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida PM70."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.
CC -1- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC KINASES.
DR EMBL: AE006176; AAK03464.1;
DR InterPro: IPR000410; Bcrl_sens.
DR InterPro: IPR000658; DUF5.
DR InterPro: IPR003660; HAMF.
DR InterPro: IPR003594; HAMF.
DR InterPro: IPR003661; His_kinA.

```

```

DR InterPro: IPR001789; Response_reg.
DR Pfam: PF00672; DUF5. 1.
DR Pfam: PF00072; response_reg. 1.
DR Pfam: PF00512; signal. 1.
DR PRINTS: PR00344; BCTRLENSOR.
DR SMART: SM00304; HAMF. 1.
DR SMART: SM00387; HAMFase_C. 1.
DR SMART: SM00388; H1SKA. 1.
DR SMART: SM00448; REC. 1.
KW Complete proteome; Kinase; Phosphorylation; Sensory transduction;
KW Transferase.
SQ SEQUENCE 862 AA; 97846 MW; AB6814A3B63626C0 CRC64;

Query Match
Best Local Similarity 2.0%; Score 9; DB 2; Length 862;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LSTLTSLT 13
DB 293 LSTLTSLT 301

RESULT 4
Q9AX85
ID 09AX85 PRELIMINARY: PRT: 1432 AA.
AC 09AX85:
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE PUTATIVE ABC TRANSPORTER PROTEIN.
GN P0410E03.7.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriophytaceae; Oryzae; Oryza.
OX NCBL_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponebare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0410E03."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP002844; BAB21276.1;
SQ SEQUENCE 1432 AA; 160234 MW; 04E8CF2CA08D993 CRC64;

Query Match
Best Local Similarity 2.0%; Score 9; DB 10; Length 1432;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 FTPLSLSD 178
DB 830 FTPLSLSD 838

RESULT 5
Q9M268
ID 09M268 PRELIMINARY: PRT: 299 AA.
AC 09M268:
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE RAV-LIKE PROTEIN.
GN P21P14.140.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OX eucotids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBL_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.

```

RA Choiane N., Robert C., Brothier P., Mincker P., Catolico L.,
RA Artigianev F., Saurin W., Weissenbach J., Mewes H.W., Lemcke K.,
RA Meyer K.F.X., Queller F., Salanoubat M.,
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL38662; CAB71904.1; -
DR InterPro: IPR003340; B3; 1.
DR Pfam: PF02362; B3; 1.
SQ SEQUENCE 299 AA; 34270 MW; 2AB841F179DD174B CRC64;

Query Match 1.8%; Score 8; DB 10; Length 299;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 174 SLSPDLDR 181
|1111111111|
DB 292 SLSPDLDR 299

RESULT 6
O9CL22 PRELIMINARY; PRT; 306 AA.
AC O9CL22;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE HYPOTHETICAL PROTEIN PM1426.
GN PM1426.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_Taxid=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=2145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L., Paustian M.L., Whittam T.S., Kapur V.,
RT "Complete genomic sequence of Pasteurella multocida pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL: AE006179; AAK03510.1; -
DR InterPro: IPR003187; PLAI.
DR Pfam: PF02253; PLAI; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 306 AA; 35580 MW; EAF3DE8C1C22B26E CRC64;

Query Match 1.8%; Score 8; DB 2; Length 306;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 314 KLSRSWNR 321
|1111111111|
DB 184 KLSRSWNR 191

RESULT 7
O67608 PRELIMINARY; PRT; 55 AA.
AC O67608;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE SUBGENOMIC DNA DERIVED FROM DNA B CCCDS - COVALENTLY CLOSED CIRCULAR
DE DOUBLE-STRANDED MOLECULE.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_Taxid=10831;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=87040767; PubMed=3022243;
RA MacDowell S.W., Coutts R.H.A., Buck K.W.,
RT "Molecular characterisation of subgenomic single-stranded and double-
RT stranded DNA forms isolated from plants infected with tomato golden
RT mosaic virus.";
RL Nucleic Acids Res. 14:7967-7984(1986).
DR EMBL: X04485; CAA28171.1; -
DR InterPro: IPR000211; Gemin1_BL.
DR Pfam: PF00845; Gemin1_BL; 1.
SQ SEQUENCE 55 AA; 6281 MW; 41963570D739EC96 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 55;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 312 SAKLSRS 318
|1111111111|
DB 43 SAKLSRS 49

RESULT 8
O91ZS5 PRELIMINARY; PRT; 72 AA.
AC O91ZS5;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY(NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_Taxid=122280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUBER NECROSING (NTN);
RA Cerovska N., Filigiarova M., Moravec T., Petrzik K.,
RT "Differences in nucleotide and amino acid sequences of N-terminal
RT parts of coat proteins among isolates of potato virus Y-NTN strain.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF228633; AAF67874.1; -
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 72 AA; 7721 MW; 2C5287E15624CBDD CRC64;

Query Match 1.6%; Score 7; DB 12; Length 72;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 156 TKDKAQ 162
|1111111111|
DB 11 TKDKAQ 17

RESULT 9
O91ZS4 PRELIMINARY; PRT; 72 AA.
AC O91ZS4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY(NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_Taxid=122280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUBER NECROSING (NTN);
RA Cerovska N., Filigiarova M., Moravec T., Petrzik K.,

RT "Differences in nucleotide and amino acid sequences of N-terminal
 parts of coat proteins among isolates of potato virus Y-NTN strain."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF228634; AAF67875.1; -
 FT NON_TER 1
 FT NON_TER 72
 SO SEQUENCE 72 AA; 7692 MW; 4F129B870523B482 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 72;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 TKKDAKQ 162
 |||||
 Db 11 TKKDAKQ 17

RESULT 10
 O91ZS3 PRELIMINARY; PRT; 95 AA.
 AC O91ZS3;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE COAT PROTEIN (FRAGMENT).
 GN CP.
 OS Potato virus Y strain NTN (PVY(NTN)).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Polyvirus.
 OX NCBI_TaxID=122280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TUBER NECROSING (NTN);
 RA Cerovska N., Filigayova M., Moravec T., Petrizik K.;
 RT "Differences in nucleotide and amino acid sequences of N-terminal
 parts of coat proteins among isolates of potato virus Y-NTN strain."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF228635; AAF67876.1; -
 FT NON_TER 1
 FT NON_TER 95
 SO SEQUENCE 95 AA; 10446 MW; 28C06891777C9928 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 95;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 TKKDAKQ 162
 |||||
 Db 11 TKKDAKQ 17

RESULT 11
 O91ZS8 PRELIMINARY; PRT; 96 AA.
 AC O91ZS8;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE COAT PROTEIN (FRAGMENT).
 GN CP.
 OS Potato virus Y (strain N) (PVY).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Polyvirus.
 OX NCBI_TaxID=12219;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NECROSING (N);
 RA Cerovska N., Filigayova M., Moravec T., Petrizik K.;
 RT "Differences in nucleotide and amino acid sequences of N-terminal
 parts of coat proteins among isolates of potato virus Y-NTN strain."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF228630; AAF67871.1; -
 FT NON_TER 1
 FT NON_TER 96
 SO SEQUENCE 96 AA; 10559 MW; 89C3A5675B9784CA CRC64;

Query Match 1.6%; Score 7; DB 12; Length 96;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 TKKDAKQ 162
 |||||
 Db 11 TKKDAKQ 17

RESULT 12
 O91ZS6 PRELIMINARY; PRT; 101 AA.
 AC O91ZS6;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE COAT PROTEIN (FRAGMENT).
 GN CP.
 OS Potato virus Y strain NTN (PVY(NTN)).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Polyvirus.
 OX NCBI_TaxID=122280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TUBER NECROSING (NTN);
 RA Cerovska N., Filigayova M., Moravec T., Petrizik K.;
 RT "Differences in nucleotide and amino acid sequences of N-terminal
 parts of coat proteins among isolates of potato virus Y-NTN strain."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF228632; AAF67873.1; -
 FT NON_TER 1
 FT NON_TER 101
 SO SEQUENCE 101 AA; 11112 MW; B5B5F51C3D5FE9BD CRC64;

Query Match 1.6%; Score 7; DB 12; Length 101;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 TKKDAKQ 162
 |||||
 Db 11 TKKDAKQ 17

RESULT 13
 O9RU16 PRELIMINARY; PRT; 122 AA.
 AC O9RU16;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE HYPOTHETICAL 13.7 KDA PROTEIN.
 GN DR1580.
 OS Deinococcus radiodurans.
 OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
 OX NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RI;
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Yamatehavan J.J., Lam P., McDonald L., Uitterback T., Zalewski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Milton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;

RT "genome sequence of the radioresistant bacterium Deinococcus
RT radiohumans R1.";
RL Science 286:1571-1577(1999).
DR EMBL: AE002002; AAF1150.1; -.
DR TIGR: DR1580; -.
DR InterPro: IPR000325; Glyoxalase_1.
DR Pfam: PF00903; Glyoxalase; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 122 AA; 13683 MW; 7BA836DCAD786E51 CRC64;

Query Match 1.6%; Score 7; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 EHEPELY 76
Db 115 EHEPELY 121

RESULT 14
Q91ZS7 PRELIMINARY; PRT; 124 AA.
AC Q91ZS7;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY(NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=122280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUBER NECROSING (NTN);
RA Cerovska N., Filigajova M., Moravec T., Petrlik K.;
RT "Differences in nucleotide and amino acid sequences of N-terminal
parts of coat proteins among isolates of potato virus Y-NTN strain.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF228631; AAF67872.1; -.
FT NON_TER 1 1
FT NON_TER 124 124
SQ SEQUENCE 124 AA; 13654 MW; DE3618D620BA7080 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 124;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 TKKDAKQ 162
Db 11 TKKDAKQ 17

RESULT 15
Q9GMD4 PRELIMINARY; PRT; 131 AA.
AC Q9GMD4;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE ALPHA-TUBULIN (FRAGMENT).
OS Macaca fuscata (Japanese macaque).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9542;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99270338; PubMed=10340513;
RA Okuno H., Tokuyama W., Li Y.X., Hashimoto T., Miyashita Y.;
RT "Quantitative evaluation of neurotrophin and trk mRNA expression in

RT visual and limbic areas along the occipito-temporo-hippocampal pathway
RT in adult macaque monkeys.";
RL J. Comp. Neurol. 408:378-398(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Hashimoto T., Tokuyama W., Okuno H., Li Y., Miyashita Y.;
RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF286378; AAC00533.1; -.
DR InterPro: IPR03008; tubulin_FtsZ.
DR Pfam: PF00091; tubulin; 1.
KW GTP-binding.
KW NON_TER 1 1
FT NON_TER 131 131
SQ SEQUENCE 131 AA; 14575 MW; A07D9C90B0641FCE CRC64;

Query Match 1.6%; Score 7; DB 6; Length 131;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120
Db 110 IKTKRSI 116

Search completed: May 9, 2002, 13:05:18
Job time: 613 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:54:58 ; Search time 72.68 Seconds
(without alignments)
450,473 Million cell updates/sec

Title: US-09-787-083-6

Perfect score: 442
Sequence: 1 MKVSLSTLTSLTSCPAIIA.....YHNEATSFVGIMLNDMGL 442

Scoring table:

OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 522463 seqs, 74073290 residues

Word size : 0

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

1: /SIDS8/gcgdata/geneseq/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	442	100.0	442	21	AAV85270
2	390	88.2	442	21	AAV85269
3	316	71.5	442	21	AAV85271
4	224	50.7	442	21	AAV85268
5	9	2.0	370	21	AAV75156
6	9	2.0	370	21	AAV75157
7	9	2.0	370	21	AAV75158
8	9	2.0	374	21	AAV70629
9	9	2.0	375	21	AAV70628
10	8	1.8	278	21	AAV1137
11	8	1.8	283	21	AAV1136

12	8	1.8	299	21	AAV1135	Arabidopsis thalia
13	7	1.6	16	15	AAV52127	Mouse light chain
14	7	1.6	33	22	AAV76059	Human colon cancer
15	7	1.6	50	21	AAV79483	Rat alpha tubulin
16	7	1.6	64	21	AAV60322	Arabidopsis thalia
17	7	1.6	82	21	AAV01488	Human secreted pro
18	7	1.6	108	21	AAV37811	Arabidopsis thalia
19	7	1.6	116	21	AAV7313	Human signal pepti
20	7	1.6	136	21	AAV37810	Arabidopsis thalia
21	7	1.6	157	21	AAV29867	Chlamydia pneumoni
22	7	1.6	169	20	AAV34661	Arabidopsis thalia
23	7	1.6	181	21	AAV37809	Envelope protein o
24	7	1.6	268	13	AAV21542	Protease inhibitor
25	7	1.6	302	13	AAV24254	Arabidopsis thalia
26	7	1.6	316	21	AAV39500	Arabidopsis thalia
27	7	1.6	324	21	AAV39499	Protease inhibitor
28	7	1.6	326	13	AAV24356	Arabidopsis thalia
29	7	1.6	365	21	AAV39498	Human endometrium
30	7	1.6	381	20	AAV60241	Human cytoskeleton
31	7	1.6	448	21	AAV1960	Arabidopsis thalia
32	7	1.6	523	21	AAV44285	Firefly luciferase
33	7	1.6	552	16	AAV67007	Amino acid sequenc
34	7	1.6	666	20	AAV21871	Amino acid sequenc
35	7	1.6	683	20	AAV21869	Amino acid sequenc
36	7	1.6	684	20	AAV21875	Amino acid sequenc
37	7	1.6	684	20	AAV21875	Amino acid sequenc
38	7	1.6	715	20	AAV21860	Amino acid sequenc
39	7	1.6	783	20	AAV70586	Human SLAP-130. H
40	7	1.6	900	20	AAV6264	P. suis leukotoxin
41	7	1.6	934	17	AAV07637	Human axlin. Homo
42	7	1.6	1026	17	AAV97374	Phage T4 tail fibr
43	7	1.6	1026	20	AAV94678	Bacteriophage T4 t
44	7	1.6	1026	20	AAV92358	Bacteriophage T4 t
45	7	1.6	1026	22	AAV35154	Bacteriophage T4 g

ALIGNMENTS

RESULT 1	
AAV85270	
ID AAV85270 standard; Protein; 442 AA.	
AC AAV85270;	
XX	
XX	
DT 29-JUN-2000 (first entry)	
XX	
XX	
DE BASB034 amino acid sequence #3.	
XX	
KW Moraxella catarrhalis infection; BASB034; diagnosis; staging;	
KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;	
KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;	
KW hearing loss; antibacterial drug.	
XX	
OS Moraxella catarrhalis.	
XX	
PN W0200015802-A1.	
XX	
PD 23-MAR-2000.	
XX	
XX	
PF 14-SEP-1999; 99WO-EP06781.	
XX	
PR 14-SEP-1998; 98GB-0020002.	
XX	
PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.	
XX	
PI Ruelle J;	
XX	
DR WPI; 2000-271440/23.	
DR N-PSDB; AAA10702.	
XX	
PT Novel BASB034 polynucleotides and polypeptides from Moraxella	
catarrhalis used to prepare vaccines against bacterial infections	

XX Claim 3; Page 68; 106pp; English.
XX
CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2913. The invention relates to BASB034 polypeptides from
CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
CC polynucleotides and polypeptides may be employed as research reagents and
CC material for the discovery of treatments and diagnostics for diseases,
CC particularly human diseases. They are particularly used to diagnose and
CC treat M. catarrhalis infections. They can be used for diagnosis of
CC disease, staging of disease, or determining response of an infectious
CC organism to drugs. The polynucleotides may be used as a source for
CC hybridization probes, and for screening of genetic mutations, serotype,
CC organism or strain identification, identification of mutations in BASB034
CC sequences, and as components of arrays which are useful for diagnostic
CC and prognostic purposes. The polypeptides can be used to produce
CC antibodies. The polypeptides can also be used in vaccine formulations,
CC and to identify agonists and antagonists. The polypeptides, antibodies,
CC agonists and antagonists (which are bacteriostatic) are used for the
CC treatment and prevention of diseases such as otitis media in infants and
CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
CC invasive diseases, and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.
XX
XX
SQ Sequence 442 AA:

Query Match 100.0%; Score 442; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKSLSTLTSLTSCFAILAIQOAKAVNPVAFVDEVRSENDLGDONELPIDVQSATQSA 60
DB 1 mkvslstltsltsclfaiaiqgkavnpvafvdevrseendlgdondelpidvqsatqsa 60
QY 61 STDTANPLDEHEPELYTTALENKTMLINCALNODIMRLACYDTLVHGETPAVTKTRST 120
DB 61 stdtanpldehepe lyttalenktmlincalnodi mrlacydtlvhgetpavtktrstl 120
QY 121 RLDETTIMQTIKGRPOVYVOETDPIFLMGNEKGMILTKKDAQOLEYAKOTPTISLSDLD 180
DB 121 rldettimqti kgrpovyvoetd piflmgnek gmltkkda qoleyakotptislsfld 180
QY 181 RNNTPPLMSRRPNPMYVLPFFMGKRPNSRPNTPSHEAROPTNPFRAPELKFOYSVYKA 240
DB 181 rnn t p l m s r r p n p m y v l p f f m g k r p n s r p n t p s h e a r o p t n p f r a p e l k f o y s v y k a 240
QY 241 AEDLMCTSDLMFGYTOQSHMQIFNCKNSRPFVNDYQPEIFLTQPVYSDLPMDGKVRMT 300
DB 241 aedlmgctsd lmfgytoqshmqifncknsrpfvndyqpeifltqpvysd lpm d g k v r m t 300
QY 301 GMGAVHHSNGESAKLSRSMNRATILMAGMEKKNLTVMPTINGRTFKESGSGOPDDNPILD 360
DB 301 gmgavhhsngesaklsrsmnratilmagmekknltvmptingrtfkessgsgopddnpild 360
QY 361 YVYGADVRFYQLEKNSNIGTVRYNPRSGKALQLDYVYPLGKIGISYQIQFGYQOSL 420
DB 361 yv yg a d v r f y q l e k n s n i g t v r y n p r s g k a l q l d y v y p l g k i g i s y q i q f g y q o s l 420
QY 421 IDYNHEATSGVGLMDMNGL 442
DB 421 idynheatsfgvglm d m n g l 442

RESULT 2
AA185269
ID AAY85269 standard; Protein; 442 AA.
XX
AC AAY85269;
XX

DT 29-JUN-2000 (first entry)
DE
XX BASB034 amino acid sequence #2.
XX
KM Moraxella catarrhalis infection: BASB034; diagnosis: staging;
KM vaccine: bacteriostatic; treatment: prevention; otitis media; pneumonia;
KM sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KM hearing loss; antibacterial drug.
XX
OS Moraxella catarrhalis.
XX
PN WO200015802-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99WO-EP06781.
XX
PR 14-SEP-1998; 98GB-0020002.
XX
PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruelle J;
XX
DR WPI: 2000-271440/23.
XX
DR N-PDB: AAA10701.
XX
PT Novel BASB034 polynucleotides and polypeptides from Moraxella
PT catarrhalis used to prepare vaccines against bacterial infections
XX
PS Claim 3; Page 67; 106pp; English.

CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2908. The invention relates to BASB034 polypeptides from
CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
CC polynucleotides and polypeptides may be employed as research reagents and
CC material for the discovery of treatments and diagnostics for diseases,
CC particularly human diseases. They are particularly used to diagnose and
CC treat M. catarrhalis infections. They can be used for diagnosis of
CC disease, staging of disease, or determining response of an infectious
CC organism to drugs. The polynucleotides may be used as a source for
CC hybridization probes, and for screening of genetic mutations, serotype,
CC organism or strain identification, identification of mutations in BASB034
CC sequences, and as components of arrays which are useful for diagnostic
CC and prognostic purposes. The polypeptides can be used to produce
CC antibodies. The polypeptides can also be used in vaccine formulations,
CC and to identify agonists and antagonists. The polypeptides, antibodies,
CC agonists and antagonists (which are bacteriostatic) are used for the
CC treatment and prevention of diseases such as otitis media in infants and
CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
CC invasive diseases, and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.
XX
XX
SQ Sequence 442 AA:

Query Match 88.2%; Score 390; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 VQASATOSASTDTANPLDEHEPELYTTALENKTMLINCALNODIMRLACYDTLVHGETPA 112
DB 53 vqatgsastdtanpldehepe lyttalenktmlincalnodi mrlacydtlvhgetp 112
QY 113 VIKTKSRIDETIMQTIKGRPOVYVOETDPIFLMGNEKGMILTKKDAQOLEYAKOTPTIS 172
DB 113 viktksr i d e t i m q t i k g r p o v y v o e t d p i f l m g n e k g m i l t k k d a q o l e y a k o t p t i s 172
QY 173 LSLSFDDLRNNTPPLMSRRPNPMYVLPFFMGKRPNSRPNTPSHEAROPTNPFRAPELK 232
DB 173 lslsfddlrnn t p l m s r r p n p m y v l p f f m g k r p n s r p n t p s h e a r o p t n p f r a p e l k f 232

QY 233 QVSVKVAADLWCTSDLMFGTQOSHWOIFNGKNSRPFRVNDYQPELITQPVYSDLP 292
CC |
CC |
CC |
Db 233 qvsvkvaadlwctsdlmfgtqgshwqifngknsrpfvndyqpeflitqpvysdlp 292
QY 293 WDGVKRMIGGAVHNSGESAKLSRSMNRAYIMAGMEKNTVMPRIWGRFRFGSGSQP 352
CC |
CC |
CC |
Db 293 wdgvkrmigmavhnsngesaklsrsmnraylmagmeknltvmpriwgrlfrfgsgsqp 352
QY 353 DDNPDILYVGYGDVRFYQLEENKSNISGTVRNPBSGKGMQLDYYVYPLGKGISGYEQI 412
CC |
CC |
CC |
Db 353 ddnpdilvygygdvrflyqlenknsisgtrvnpbsgkqmlddyvypylgkgisgyfqi 412
QY 413 FQGTGQSLIDYVNHGATSFVGLMLNDWACL 442
CC |
CC |
CC |
Db 413 fqgtgqslldynheatsfygqlmndwmg1 442

RESULT 3
AAV85271
ID AAV85271 standard; Protein: 442 AA.
XX
XX AAV85271;
XX
XX 29-JUN-2000 (first entry)
XX
XX BASB034 amino acid sequence #4.
XX
XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KM vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KM sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KM hearing loss; antibacterial drug.
XX
XX Moraxella catarrhalis.
XX OS
XX WC200015802-A1.
XX
XX 23-MAR-2000.
XX
XX 14-SEP-1999; 99WG-EP06781.
XX
XX 14-SEP-1998; 98GB-0020002.
XX
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Ruelle J;
XX
XX WPI: 2000-271440/23.
XX
XX N-PSDB: AAA10703.

Novel BASB034 polynucleotides and polypeptides from Moraxella
catarrhalis used to prepare vaccines against bacterial infections -
PS Claim 3; Page 69; 106pp; English.

This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
strain Mc2969. The invention relates to BASB034 polypeptides from
M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
polynucleotides and polypeptides may be employed as research reagents and
material for the discovery of treatments and diagnostics for diseases,
particularly human diseases. They are particularly used to diagnose and
treat M. catarrhalis infections. They can be used for diagnosis of
disease, staging of disease, or determining response of an infectious
organism to drugs. The polynucleotides may be used as a source for
hybridization probes, and for screening of genetic mutations, serotype,
organism or strain identification, identification of mutations in BASB034
sequences, and as components of arrays which are useful for diagnostic
and prognostic purposes. The polypeptides can be used to produce
antibodies. The polypeptides can also be used in vaccine formulations,
and to identify agonists and antagonists. The polypeptides, antibodies,
agonists and antagonists (which are bacteriostatic) are used for the
treatment and prevention of diseases such as otitis media in infants and
children, pneumonia in elderlies, sinusitis, nosocomial infections and

CC invasive diseases, and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.
XX
SQ Sequence 442 AA;
Query Match 71.5%; Score 316; DB 21; Length 442;
Best Local Similarity 99.8%; Pred. No. 3.6e-298;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 26 AVNPVAFVDEVRSENDLGQDNELPIDYQASQASITDANLDEHEPELTTALENKT 85
CC |
CC |
CC |
Db 26 avnpvafvdevrsendlgqdnelpidyqasqasitdanldehepeltytalenktm 85
QY 86 LINCASALNODIMRLACYDTLVHGFPAVITKRSIRIDETIMOTIKGPOVVOETDP 145
CC |
CC |
CC |
Db 86 lincsalnqdimrlacydclvhgfepavltkrsiridetimotikgppvyvqetcp1 145
QY 146 FLMGNEKGMILTKDAKOLEYAAKQFTPLSFDLRNNTPLMSSRPHPMTVLEPFM 205
CC |
CC |
CC |
Db 146 flmgnekgmiltkdakqleyaakqftplsfldrnnptlmsrphpmvylepfmngk 205
QY 206 PNRSPNTPSHEARQFTPEFRAPBLKFOVSVKVAADLWCTSDLMFGTQOSHWOIFN 265
CC |
CC |
CC |
Db 206 pnrspntpshearqftpefrapblkfgyvsvkvaadlwctsdlmfgtqgshwq1fn 265
QY 266 GKNSRPRVHNOYQPELITQPVYSDLPMDGKVRMIGGAVHNSGESAKLSRSMNRAYIM 325
CC |
CC |
CC |
Db 266 gknsrprvhnoyqpeflitqpvysdlpmdgkvrmgigmavhnsngesaklsrsmnray1m 325
QY 326 AGMEKMLTVMPRIWGRIFKESGSGQPDNPDILYVGYGDVRFYQLEENKSNISGTVRY 385
CC |
CC |
CC |
Db 326 agmekmltvmprlwgtrifksgsgqpdnpdildygygdvrflyqlenknsisgtrvy 385
QY 386 NPSRGKALQLDYVYPPLAGKISGTFQGTGQSLIDYVNHGATSFVGLMLNDWACL 442
CC |
CC |
CC |
Db 386 npsrgkalldyvypplyagkisgtrfqgtgqslldynheatsfygqlmndwmg1 442

RESULT 4
AAV85268
ID AAV85268 standard; Protein: 442 AA.
XX
XX AAV85268;
XX
XX 29-JUN-2000 (first entry)
XX
XX BASB034 amino acid sequence #1.
XX
XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KM vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KM sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KM hearing loss; antibacterial drug.
XX
XX Moraxella catarrhalis.
XX OS
XX WC200015802-A1.
XX
XX 23-MAR-2000.
XX
XX 14-SEP-1999; 99WG-EP06781.
XX
XX 14-SEP-1998; 98GB-0020002.
XX
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Ruelle J;
XX
XX WPI: 2000-271440/23.
XX
XX N-PSDB: AAA10700.


```

PF 30-APR-1999; 99WO-US09346.
XX
PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
PA (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
DR WPI: 2000-062150/05.
DR N-PSDB; AA253919.
XX
PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics
XX
PS Claim 2; Page 904; 1453pp; English.
XX
CC AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
CC and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 370 AA;

Query Match 2.0%; Score 9; DB 21; Length 370;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIFLTQPV 287
Db 209 peifltqpv 217

RESULT 7
AAY75158
ID AAY75158 standard; Protein; 370 AA.
XX
AC AAY75158;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 582 protein sequence SEQ ID NO:1790.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
KW antibacterial; gene therapy.
XX
OS Neisseria meningitidis.
XX
PN WO957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.

```

```

XX
PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
PA (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
DR WPI: 2000-062150/05.
DR N-PSDB; AA253920.
XX
PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics
XX
PS Claim 2; Page 905; 1453pp; English.
XX
CC AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
CC and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 370 AA;

Query Match 2.0%; Score 9; DB 21; Length 370;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIFLTQPV 287
Db 209 peifltqpv 217

RESULT 8
AAY70629
ID AAY70629 standard; Protein; 374 AA.
XX
AC AAY70629;
XX
DT 18-JUL-2000 (first entry)
XX
DE Neisseria meningitidis serogroup B strain H44/76 BASB033 protein.
XX
KW BASB033; diagnosis; prophylaxis; treatment; antibacterial; vaccine;
KW Neisseria meningitidis infection.
XX
OS Neisseria meningitidis.
XX
PN WO200015801-A1.
XX
PD 23-MAR-2000.
XX
PF 09-SEP-1999; 99WO-EP06718.
XX
PR 14-SEP-1998; 98GB-0020003.

```

```

XX PA      (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX PI
XX Ruelle J;
DR MPI: 2000-271439/23.
XX N-PSTB: AA6252134.
PT Isolated BASB033 polypeptides and polynucleotides of Neisseria meningitidis, useful for diagnosis, prophylaxis and treatment of N. meningitidis infection -
PS Claim 4 : Page 59; 93pp; English.
XX CC The present sequence is a BASB033 protein from
CC Nelseteria meningitidis serogroup B strain H44/76. The protein shows homology to the Klebsiella pneumoniae outer membrane phospholipase A. The present sequence is useful for diagnosis, prophylaxis and treatment of N. meningitidis infection. It may also be used for the discovery and development of antibacterial compounds and in vaccine compositions.
XX SQ Sequence 374 AA;
OY Query Match 2.0%; Score 9; DB 21; Length 374;
Best Local Similarity 100.0%; Pred.No.2.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 213 Peifilpv 221
IIIIIIIII
Peifilpvy 287

RESULT          9
AAV70628
ID AAV70628 standard; Protein: 375 AA.
AC AAY70628;
AT 18-JUL-2000 (first entry)
DE Neisseria meningitidis serogroup B strain ATCC13090 BASB033 protein.
KW BASB033; diagnosis; prophylaxis; treatment; antibacterial; vaccine;
RV Neisseria meningitidis infection.
OS Neisseria meningitidis.
XX WO200015801-A1.
PD 23-MAR-2000.
XX PF 09-SEP-1999; 99WO-EP06718.
PR 14-SEP-1998; 98GB-0020003.
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
PA Ruelle J;
PI
XX DR MPI: 2000-271439/23.
XX N-PSTB: AA6252133.
PT Isolated BASB033 polypeptides and polynucleotides of Neisseria meningitidis, useful for diagnosis, prophylaxis and treatment of N. meningitidis infection -
PS Claim 4 : Page 58; 93pp; English.
XX CC The present sequence is a BASB033 protein from
CC Neisseria meningitidis serogroup B strain ATCC13090. The protein shows homology to the Klebsiella pneumoniae outer membrane

```

CC	phospholipase A. The present sequence is useful for diagnosis,
CC	prophylaxis and treatment of N. meningitidis infection. It may also be
CC	used for the discovery and development of antibacterial compounds and
CC	in vaccine compositions.
XX	
SQ	Sequence 375 AA;
Oy	279 PEIPLTOPY 287
Db	214 peifltgpy 222
RESULT 10	
AAAG1137	
ID	AAAG1137 standard; Protein: 278 AA.
XX	
AC	AAAG1137;
DT	17-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 9741.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	
XX	Arabidopsis thaliana.
OS	
PM	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PE	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130510.
PR	28-APR-1999; 99US-0130891.
PR	30-APR-1999; 99US-0131449.
PR	30-APR-1999; 99US-0132048.
PR	04-MAY-1999; 99US-0132407.
PR	04-MAY-1999; 99US-0132484.
PR	05-MAY-1999; 99US-0132485.
PR	06-MAY-1999; 99US-0132486.
PR	07-MAY-1999; 99US-0132487.
PR	11-MAY-1999; 99US-0132863.
PR	11-MAY-1999; 99US-0134256.
PR	14-MAY-1999; 99US-0134218.
PR	14-MAY-1999; 99US-0134219.
PR	14-MAY-1999; 99US-0134221.
PR	14-MAY-1999; 99US-0134370.
PR	18-MAY-1999; 99US-0134768.
PR	19-MAY-1999; 99US-0134941.
PR	20-MAY-1999; 99US-0135124.
PR	21-MAY-1999; 99US-0135353.
PR	24-MAY-1999; 99US-0135629.
PR	25-MAY-1999; 99US-0136021.
PR	27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159350.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 8; DB 21; Length 283;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 174 SLSEFLDR 181
|||||||

DB 276 slsflidr 283
RESULT 12
AAG1135
ID AAG1135 standard; Protein: 299 AA.
XX
AC AAG1135;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 9739.
XX
KW Protein identification: signal transduction pathway; metabolic pathway;
KW hydrolisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127452.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137232.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139763.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141642.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144684.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145226.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146388.
PR 03-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.

PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149992.
PR 23-AUG-1999; 9905-0149992.
PR 23-AUG-1999; 9905-0149993.
PR 23-AUG-1999; 9905-0149993.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151330.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158332.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 25-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161932.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

Query Match 1.88; Score 8; DB 21; Length 299;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 SLSPDLR 181
Db 292 slsrdlar 299

RESULT 13
AAK52127
ID AAK52127 standard; Peptide; 16 AA.
XX

AC AAR52127;
 XX
 DT 27-SEP-1996 (first entry)
 XX
 DE Mouse light chain surface patch 503410.
 XX
 KW antibody: humanised; murine; human; heavy chain; light; variable;
 KW framework region; complementarity determining region; reshaping;
 KW modelling; surface residue; modify.
 XX
 OS Mus sp.
 XX
 PN EP592106-A1.
 XX
 PD 13-APR-1994.
 XX
 PE 07-SEP-1993; 93EP-0307051.
 XX
 PR 09-SEP-1992; 92US-0942245.
 XX
 PA (PEDE/) PEDERSEN J T.
 PA (IMMU-) IMMUNOGEN INC.
 XX
 PI Guild BC, Pedersen JT, Rees AR, Roguska MA, Searle SMJ;
 XX
 DR WPI; 1994-120230/15.
 XX
 PT Method of resurfacing of rodent antibodies to produce humanised
 PT antibody forms - for producing non-human antibodies with improved
 PT therapeutic efficiency by presenting human surface on V-region
 XX
 PS Example 1; Page 14; 230pp; English.
 XX
 CC Modification of a rodent antibody or fragment by resurfacing in order
 CC to produce a humanised rodent antibody can be determined by calculating
 CC homology between murine and human antibody surfaces. In order to test
 CC the resurfacing approach of the invention, three humanisation
 CC experiments were set up. (1) traditional loop grafting; (2) resurfacing
 CC approach using most similar chain; and (3) resurfacing approach using
 CC human sequences with most similar surface residues. AAR52069-159 are the
 CC surface residue patterns in mouse light chain antibody variable regions.
 CC These "patches" were used in the third method, where rodent light and
 CC heavy chains were matched and the most similar human sequence found
 CC independently only over the surface residues indicated in AAR52030-67.
 XX
 SQ Sequence 16 AA:
 OY
 DB 387 PRSGKA 393
 111111
 4 prsgka 10
 Query Match 1.6%; Score 7; DB 15; Length 16;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 14
 AAG76069
 ID AAG76069 standard; Protein: 33 AA.
 AC AAG76069;
 XX
 DT 03-SEP-2001 (first entry)
 XX
 DE Human colon cancer antigen protein SEQ ID NO:6833.
 XX
 KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma.
 XX
 OS Homo sapiens.
 XX
 PN WO200122920-A2.

XX
 PD 05-APR-2001.
 XX
 XX 28-SEP-2000; 2000WO-US26524.
 XX
 PE 29-SEP-1999; 99US-0157137.
 XX
 PR 03-NOV-1999; 99US-0163280.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Barash SC, Birse CE, Rosen CA;
 XX
 DR WPI: 2001-235357/24.
 DR N-PSDB: AAR35474.
 XX
 PT Nucleic acids encoding 427 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -
 XX
 PS Claim 11; Page 8285-8286; 9803pp; English.
 XX
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patients own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated Ps,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX
 SQ Sequence 33 AA:
 OY
 DB 8 LTLSTLS 14
 111111
 22 ltlstls 28
 Query Match 1.6%; Score 7; DB 22; Length 33;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 15
 AAY79483
 ID AAY79483 standard; peptide: 50 AA.
 AC AAY79483;
 XX
 DT 01-AUG-2000 (first entry)
 XX
 DE Rat alpha tubulin variant peptide (aa314-363).
 XX
 KW Rat; alpha tubulin; mutant; mutein; variant;
 KW metabolic x syndrome; hypertension; stroke; diabetes;
 KW insulin resistance; obesity; dyslipidemia; anorectic;
 KW hypotensive; cerebroprotective; vasotropic; antidiabetic;
 KW antilipemic; gene therapy; spontaneously hypertensive; SHR.
 XX
 OS Rattus sp.
 XX
 FH Key
 FT Peptide
 Location/Qualifiers
 1..50
 /note="corresponds to residues 314-363 of

```

FT      Misc-difference 27      full-length protein"
FT      /note- "replaces Thr in wild-type"
XX
XX      MO200018918-A2.
XX
XX      06-APR-2000.
XX
XX      28-SEP-1999; 99MO-US22494.
XX
XX      28-SEP-1998; 98US-0161939.
XX
XX      (CURA-) CURAGEN CORP.
XX
XX      Shimkets RA;
XX
XX      WPI; 2000-303450/26.
XX
XX      Novel genes encoding e.g. CD36, SGLT2, and kynurenine aminotransferase,
PT      which are predictive and therapeutic for stroke, hypertension, diabetes
PT      and obesity -
XX
XX      Claim 2; Fig 4G; 79pp; English.
XX
XX      This sequence represents amino acids 314-363 of alpha-tubulin
CC      from a spontaneously hypertensive rat (SHR). This region of
CC      alpha-tubulin carries a Ser for Thr amino acid substitution when
CC      compared to the corresponding region in control Wistar Kyoto (WKY)
CC      rats (see MAY79484). The alpha-tubulin gene is differentially
CC      expressed in SHR, SHR-stroke prone (SP) and WKY rats. It may be
CC      involved in a predisposition to vascular injury. The SHR rat
CC      is an animal model for human metabolic X syndrome. Genes encoding
CC      sodium dependent glucose cotransporter, kynurenine aminotransferase,
CC      CD36, aldolase A, atrial natriuretic factor, alpha-cardiac myosin
CC      and alpha-tubulin were identified as being potentially associated
CC      with hypertension, obesity and insulin resistance. These
CC      proteins, including isolated proteins comprising the present
CC      sequence, are used for treating, preventing and diagnosing
CC      ischemic and metabolic diseases and disorders, such as stroke,
CC      hypertension, diabetes and obesity, especially insulin resistivity,
CC      dyslipidemia and ischemic stroke (all claimed). Polynucleotides
CC      encoding them may also be used in gene therapy and antisense
CC      therapy protocols.
XX
XX      Sequence 50 AA:
SQ

```

Query Match 1.6%; Score 7; DB 21; Length 50;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      114 IKTIRST 120
        |||||
Db       22 Iktkrs1 28

```

Search completed: May 9, 2002, 12:54:59
Job time: 514 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:55:42 ; Search time 36.87 seconds

(without alignments)
269.771 Million cell updates/sec

Title: US-09-787-083-6

Perfect score: 442
Sequence: 1 MKVSLSTLSTLSCFALLA.....YNHATSPGVGLMDNMGL 442

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 212252 seqs, 22503292 residues

Word size : 0

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA: *
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep: *
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/PCFUS_COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	1.6	16	1	US-07-942-245-98
2	7	1.6	347	4	US-09-097-889-14
3	7	1.6	552	1	US-08-231-7298-6
4	7	1.6	900	4	US-08-890-865A-4
5	7	1.6	934	1	US-08-215-805A-80
6	7	1.6	1026	2	US-08-542-003-6
7	7	1.6	1026	2	US-08-322-760A-6
8	6	1.4	10	3	US-08-974-775-30
9	6	1.4	11	3	US-08-974-775-29
10	6	1.4	12	3	US-08-974-775-28
11	6	1.4	13	3	US-08-974-775-8
12	6	1.4	13	3	US-08-974-775-27
13	6	1.4	14	3	US-08-974-775-9
14	6	1.4	14	3	US-08-974-775-10
15	6	1.4	14	3	US-08-974-775-12
16	6	1.4	15	3	US-08-974-775-5
17	6	1.4	16	3	US-08-974-775-4
18	6	1.4	20	1	US-08-430-273-1
19	6	1.4	21	1	US-07-918-953-13
20	6	1.4	21	1	US-07-918-953-15
21	6	1.4	21	1	US-08-212-696-1
22	6	1.4	21	1	US-08-158-245-1
23	6	1.4	21	1	US-08-081-661-13
24	6	1.4	21	1	US-08-081-661-15
25	6	1.4	21	1	US-08-233-617-1
26	6	1.4	21	1	US-08-160-376A-1
27	6	1.4	21	1	US-08-304-070-1

28	6	1.4	21	1	US-08-285-661-1	Sequence 1, Appl
29	6	1.4	21	1	US-08-301-838-1	Sequence 1, Appl
30	6	1.4	21	1	US-08-389-487-4	Sequence 4, Appl
31	6	1.4	21	1	US-08-389-487-12	Sequence 12, Appl
32	6	1.4	21	1	US-08-507-124-2	Sequence 2, Appl
33	6	1.4	21	1	US-08-507-124-3	Sequence 3, Appl
34	6	1.4	21	1	US-08-342-931-1	Sequence 1, Appl
35	6	1.4	21	1	US-08-400-256-1	Sequence 1, Appl
36	6	1.4	21	2	US-08-508-664-9	Sequence 9, Appl
37	6	1.4	21	2	US-08-353-476-85	Sequence 85, Appl
38	6	1.4	21	2	US-08-353-476-87	Sequence 87, Appl
39	6	1.4	21	2	US-08-484-219-1	Sequence 1, Appl
40	6	1.4	21	2	US-08-979-587-1	Sequence 1, Appl
41	6	1.4	21	2	US-08-992-676-1	Sequence 6, Appl
42	6	1.4	21	2	US-08-992-676-6	Sequence 1, Appl
43	6	1.4	21	2	US-09-134-836-1	Sequence 1, Appl
44	6	1.4	21	2	US-09-134-836-7	Sequence 7, Appl
45	6	1.4	21	4	US-08-932-082-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-07-942-245-98
Sequence 98, Application US/07942245
Patent No. 5639641
GENERAL INFORMATION:
APPLICANT: PEDERSEN, Jan T.
APPLICANT: SEARLE, Stephen M.J.
APPLICANT: REES, Anthony R.
APPLICANT: ROGUSKA, Michael A.
APPLICANT: GUILD, Brydon C.
TITLE OF INVENTION: SURFACE RESIDUE VENERING OF RODENT
NUMBER OF SEQUENCES: 522
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughruue, Mion, Zim, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3302
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: HP 9000/700 Workstation
OPERATING SYSTEM: UNIX
SOFTWARE: In house
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/942,245
FILING DATE: 09-SEP-1992
CLASSIFICATION: 530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 98:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-942-245-98

Query Match 1.6%; Score 7; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 7; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
QY 387 PRSGGA 393
|||||
DB 4 PRSGGA 10

RESULT 2
US-09-097-889-14
; Sequence 14, Application US/09097889
; Patent No. 6218117
; GENERAL INFORMATION:
; APPLICANT: Herrstadt, Corrina
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Davis, Robert E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
; TITLE OF INVENTION: AGENTS THAT QUANTITATIVELY ALTER DETECTABLE
; TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA: MITOCHONDRIAL DNA RATIOS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/097,889
; APPLICATION NUMBER: US/09/097,889
; FILING DATE: 15-JUN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Roseman Ph.D., Stephen J.
; REGISTRATION NUMBER: 43,058
; REFERENCE/DOCKET NUMBER: 660088.417
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ. ID NO.: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-09-097-889-14

Query Match 1.6%; Score 7; DB 4; Length 347;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 LTLSTLS 14
|||||
Db 155 LTLSTLS 161

RESULT 3
US-08-231-729B-6
; Sequence 6, Application US/08231729B
; Patent No. 5618722
; GENERAL INFORMATION:
; APPLICANT: ZENNO, Shubel
; APPLICANT: SHIRAIISHI, Shinji
; APPLICANT: INOUE, Satoshi
; APPLICANT: SAIGO, Kaoru
; TITLE OF INVENTION: FIREFLY LUCIFERASE GENE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEYDIG, VOIT & MAYER
; STREET: 700 Thirteenth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,729B
; FILING DATE: 20-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 119050/1993
; FILING DATE: 21-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rose, Herbert C.
; REGISTRATION NUMBER: 29846
; REFERENCE/DOCKET NUMBER: 60130/No. 5618722aka
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-737-6770
; TELEFAX: 202-737-6776
; INFORMATION FOR SEQ. ID NO.: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 552 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-231-729B-6

Query Match 1.6%; Score 7; DB 1; Length 552;
Best Local Similarity 100.0%; Pred. No. 11e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 175 LSFDLDR 181
|||||
Db 184 LSFDLDR 190

RESULT 4
US-08-890-865A-4
; Sequence 4, Application US/08890865A
; Patent No. 6307019
; GENERAL INFORMATION:
; APPLICANT: Constantin, Franklin
; TITLE OF INVENTION: AXIN GENE AND USES THEREOF
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/890,865A
; FILING DATE: 10-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/54249
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)278-0400
; TELEFAX: (212)391-0526
; INFORMATION FOR SEQ. ID NO.: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 900 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-890-865A-4

Query Match 1.6%; Score 7; DB 4; Length 900;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 403 GKISGY 409
|||||
DB 268 GKISGY 274

RESULT 5
US-08-215-805A-80
Sequence 80, Application US/08215805A
Patent No. 5559008
GENERAL INFORMATION:
APPLICANT: Chang, Yung-Fu
TITLE OF INVENTION: LEUKOTOXIN GENE FROM PASTEURRELLA
TITLE OF INVENTION: SUI5
NUMBER OF SEQUENCES: 84
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/215,805A
FILING DATE: 22-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Timlan, Susan J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 19603/61 (D-1329A)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1636
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 934 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Pasteurella suis
STRAIN: 5943
IMMEDIATE SOURCE:
LIBRARY: P. suis DNA in Bacteriophage lambda-dash
CLONE: (Lambda)yfc33-37
US-08-215-805A-80

Query Match 1.6%; Score 7; DB 1; Length 934;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 238 VKAEDL 244
|||||
DB 62 VKAEDL 68

RESULT 6
US-08-542-003-6

Sequence 6, Application US/08542003
Patent No. 5864013
GENERAL INFORMATION:
APPLICANT: Goldberg, Edward B.
TITLE OF INVENTION: MATERIALS FOR THE PRODUCTION OF
TITLE OF INVENTION: NANOMETER STRUCTURES AND USE THEREOF
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie and Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/542,003
FILING DATE: 13-OCT-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 8471-0005-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: 212-869-8664
TELEX: 66441 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1026 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Bacteriophage T4
IMMEDIATE SOURCE:
CLONE: p37 amino acid
US-08-542-003-6

Query Match 1.6%; Score 7; DB 2; Length 1026;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 128 QTIKGP 134
|||||
DB 861 QTIKGP 867

RESULT 7
US-08-322-760A-6
Sequence 6, Application US/08322760A
Patent No. 5877279
GENERAL INFORMATION:
APPLICANT: Goldberg, Edward B.
TITLE OF INVENTION: MATERIALS FOR THE PRODUCTION OF
TITLE OF INVENTION: NANOMETER STRUCTURES AND USE THEREOF
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie and Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/322,760A
FILING DATE: 13-OCT-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mastrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 8471-0003-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: 212-869-8864
TELEX: 66441 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1026 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Bacteriophage T4
IMMEDIATE SOURCE:
CLONE: p37 amino acid
US-08-322-760A-6

Query Match 1.6%; Score 7; DB 2; Length 1026;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 QTIKGP 134
|||||||
DB 861 QTIKGP 867

RESULT 8
US-08-974-775-30
Sequence 30, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
APPLICANT: Toback, F. Gary
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
STREET: Plaza Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775
FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid

STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-30

Query Match 1.4%; Score 6; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
|||||||
DB 5 NHEATS 10

RESULT 9
US-08-974-775-29
Sequence 29, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
APPLICANT: Toback, F. Gary
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
STREET: Plaza Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775
FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-29

Query Match 1.4%; Score 6; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
|||||||
DB 5 NHEATS 10

RESULT 10
US-08-974-775-28
Sequence 28, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:

APPLICANT: Toback, F. Gary
APPLICANT: Walsh-Reitz, Margaret
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
STREET: Plaza Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775
FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-28

Query Match 1.4%; Score 6; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 424 NHEATS 429
Db 5 NHEATS 10
RESULT 11
US-08-974-775-8
Sequence 8, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
APPLICANT: Toback, F. Gary
APPLICANT: Walsh-Reitz, Margaret
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
STREET: Plaza Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775

FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-8

Query Match 1.4%; Score 6; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 424 NHEATS 429
Db 8 NHEATS 13

RESULT 12
US-08-974-775-27
Sequence 27, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
APPLICANT: Toback, F. Gary
APPLICANT: Walsh-Reitz, Margaret
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
STREET: Plaza Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775
FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-27

Query Match 1.4%; Score 6; DB 3; Length 13;

Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 424 NHEATS 429
|||||
DB 5 NHEATS 10

RESULT 13
US-08-974-775-9
; Sequence 9, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reltz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; STREET: Plaza Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ. ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-974-775-9

Query Match 1.4%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 424 NHEATS 429
|||||
DB 8 NHEATS 13

RESULT 14
US-08-974-775-10
; Sequence 10, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reltz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE

; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; STREET: Plaza Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ. ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-974-775-10

Query Match 1.4%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 424 NHEATS 429
|||||
DB 8 NHEATS 13

RESULT 15
US-08-974-775-32
; Sequence 32, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reltz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; STREET: Plaza Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-321-4200
 TELEFAX: 312-321-4299
 INFORMATION FOR SEQ ID NO: 32:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 14 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FEATURE:
 NAME/KEY: modified-site
 LOCATION: 14
 OTHER INFORMATION: /product= "Ser or Tyr"
 US-08-974-775-32

Query Match 1.4%; Score 6; DB 3; Length 14;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 424 NHEATS 429
 |||||
 Db 8 NHEATS 13

Search completed: May 9, 2002, 12:55:42
 Job time: 552 sec

This Page Blank (uspto)